

Secrecy from Resolvability

Matthieu R. Bloch and J. Nicholas Laneman

Abstract

We investigate an approach to physical-layer security based on the premise that the coding mechanism for secrecy over noisy channels is fundamentally tied to the notion of resolvability. Instead of considering *capacity-based* constructions, which associate to each message a sub-code whose rate approaches the capacity of the eavesdropper's channel, we consider *resolvability-based* constructions, which associate to each message a sub-code whose rate is beyond the resolvability of the eavesdropper's channel. We provide evidence that resolvability is a more powerful and perhaps more fundamental coding mechanism for secrecy by developing results that hold for strong secrecy metrics and arbitrary channels. Specifically, we show that, at least for binary symmetric wiretap channels, random capacity-based constructions fail to approach the strong secrecy capacity while resolvability-based constructions achieve it. We then obtain the secrecy-capacity region of arbitrary broadcast channels with confidential messages and a cost constraint for strong secrecy metrics, which generalizes existing results. Finally, we specialize our results to study the secrecy capacity of wireless channels with perfect channel state information, compound and mixed channels, as well as the secret-key capacity of source models for secret-key agreement. By tying secrecy to resolvability, we obtain achievable rates for stronger secrecy metrics and with simpler proofs than previously derived.

Index Terms

information-theoretic security, wiretap channel, secret-key agreement, information-spectrum, resolvability, wireless channels.

Matthieu R. Bloch is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, and with the GT-CNRS UMI 2968, Metz, France. J. Nicholas Laneman is with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN. Parts of these results were presented at the 46th Allerton Conference on Communication, Controls, and Computing, Monticello, IL and at the 2011 IEEE International Symposium on Information Theory, Saint-Petersburg, Russia.

I. INTRODUCTION

In virtually every communication system, the problems of reliability and secrecy are handled in fundamentally different ways. Typically, error-correcting schemes in the physical-layer guarantee reliable communications, while encryption algorithms and key-exchange protocols in the upper layers¹ ensure data secrecy. Physical-layer security puts forward an alternative role for the physical layer, whereby reliability and secrecy can be handled jointly by means of appropriate coding schemes. The premise of physical-layer security is to recognize the presence of noise in every communication channel, including the channel of a potential adversary who eavesdrops on transmitted signals, and to exploit knowledge of noise statistics to prevent eavesdroppers from retrieving information. Unlike usual security schemes, physical-layer security can guarantee information-theoretic security, by which secrecy is measured quantitatively in terms of the statistical independence between the messages transmitted and the observations of eavesdroppers.

The theoretical foundations of physical-layer security build upon the early works of Wyner [1] and Csiszár & Körner [2], which prove the existence of coding schemes ensuring reliability and secrecy for the wiretap channel; however, the recent surge of information-theoretic results about the wiretap channel has fostered few practical engineering solutions. This state of affairs is partly due to the fact that most works exploit the coding schemes of [1], [2], in which the coding mechanism that guarantees secrecy is tied to channel capacity. This mechanism will be precisely defined in Section III; at this point, suffice to say that the codes in [1], [2] are a union of sub-codes whose rates approach the channel capacity of the eavesdropper's channel as the blocklength grows large. Although such coding schemes have been successfully used to study many multiuser information-theoretic secrecy problems [3], [4], deriving secrecy from channel capacity leaves open a few lingering issues:

- 1) wiretap channel models that incorporate the limitations of modern communication systems, such as memory or lack of channel state knowledge, are difficult to analyze;
- 2) the results obtained by tying secrecy to channel capacity are deemed too weak for cryptographic applications.

This paper discusses an alternative approach to physical-layer security that addresses the aforementioned issues; the premise of the approach is that the coding mechanism for secrecy is fundamentally related to the notion of resolvability [5] and not to channel capacity.

¹Specific cryptographic schemes are implemented at all upper layers of the protocol stack, including MAC, transport, network, and application layers.

A. Motivating Examples

To motivate the approach, we start with two intuitive examples that shed light on the mechanisms one could exploit to ensure information-theoretic security.

Example 1 (One-time pad). Consider a binary message $M \in \{0, 1\}$ that is encoded into a codeword Z as $Z = M \oplus K$, where $K \sim \mathcal{B}(p)$ is a secret key and \oplus denotes the modulo-two addition. If $p = \frac{1}{2}$, the crypto lemma [6] shows that the output distributions $p_{Z|M=0}$ and $p_{Z|M=1}$ are identical and equal to the uniform distribution on $\{0, 1\}$; hence, messages are statistically indistinguishable for an eavesdropper observing Z alone. From an operational perspective, note that the encoder exploits the key K to ensure that all messages induce the same output distribution.

Example 2 (Transmission over a noisy Gaussian channel). Consider an uncoded message M uniformly distributed in the set $\{-1, +1\}$ and observed by an eavesdropper at the output of a real additive white Gaussian noise channel as $Z = M + N$, where $N \sim \mathcal{N}(0, \sigma^2)$. As illustrated in Figure 1, the output distributions $p_{Z|M=-1}$ and $p_{Z|M=+1}$ become indistinguishable from the average distribution p_Z as the noise variance increases. Specifically, as σ goes to infinity, one can show that, for each $m \in \{-1; +1\}$,

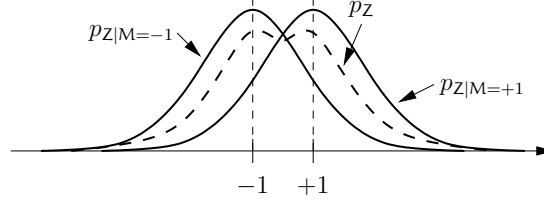


Fig. 1. Distributions of channel outputs over AWGN channel.

the variational distance between $p_{Z|M=m}$ and p_Z satisfies

$$\int_{\mathbb{R}} |p_{Z|M=m}(z) - p_Z(z)| dz = \mathcal{O}(\sigma^{-\frac{3}{2}}).$$

In other words, if the channel introduces enough randomness, then the channel itself ensures that all messages induce approximately the same output distribution.

In both examples, statistical indistinguishability is obtained because there exists a source of randomness (key or channel noise) and a coding mechanism by which all messages induce the same distribution for the eavesdropper's observations; this operation is reminiscent of the codes analyzed in [5], [7] to study the notion of *resolvability*. At this point, the connection between secrecy and resolvability may seem

contrived but, nevertheless, it suggests the possibility of ensuring secrecy by means that are radically different from those based on channel capacity and used in [1], [2]. In the remainder of this paper, we develop a set of results that expand upon the ideas introduced in Example 1 and Example 2. We not only highlight the benefits of explicitly connecting secrecy to resolvability but also show the limitations of an approach based on channel capacity.

B. Related Work

Most communication architectures providing information-theoretic security are based on two models of communication. The *wiretap channel*, introduced by Wyner [1] and generalized by Csiszár & Körner [2], models an architecture in which a transmitter encodes messages M into codewords X^n of n symbols for transmission to a receiver, in the presence of an eavesdropper who obtains noisy observations Z^n of X^n . In the case of discrete memoryless channels, [1], [2] have shown the existence of coding schemes simultaneously ensuring reliable transmission to the receiver and secrecy with respect to the eavesdropper. In particular, it is possible to characterize the *secrecy capacity* of a wiretap channel, defined as the supremum of all reliable and secure rates. The extension of this result to Gaussian [8] and wireless channels (see, for instance, [9] and references therein) suggests the potential of such coding schemes to secure communication networks at the physical layer. An alternative to the wiretap channel is the *source model for secret-key agreement* introduced by Maurer [10] and Ahlswede & Csiszár [11], which considers an architecture in which two legitimate parties attempt to distill secret keys from a noisy source by communicating over a public channel. The resulting keys have to be secure with respect to an eavesdropper who obtains correlated observations from the source and observes all messages exchanged over the public channel. This architecture differs from the wiretap channel by exclusively focusing on the rate of secret key that can be distilled from the source and by ignoring the cost of public communication. The counterpart of secrecy capacity is the *secret-key capacity*, defined as the supremum rate of secret keys that can be distilled. Although the aforementioned architectures model fundamentally different communication scenarios, they are related in that a coding scheme for the wiretap channel can be used to design a coding scheme for secret-key agreement and vice-versa.

The information-theoretic security results obtained for the wiretap channel and source model for secret-key agreement are criticized in some circles for measuring statistical independence in terms of the rate of information leaked to the eavesdropper $\frac{1}{n} \mathbb{I}(M; Z^n)$. The weakness of this metric from a cryptographic standpoint has been highlighted in multiple works [4], [12], which have advocated using the total amount of information leaked $\mathbb{I}(M; Z^n)$ instead. The analysis of secure communication architectures

under this more stringent secrecy metric has been performed with different methods, such as graph-coloring techniques [13] and privacy amplification [12], [14]. We also note that resolvability has already been used more or less implicitly in [15], [16]; the results presented in this paper differ from these earlier works by making resolvability the *explicit* mechanism for secrecy and generalizing known results to several models, including compound channels and continuous channels with cost constraints.

The connection between secrecy and resolvability is better highlighted by studying secure communication architectures beyond the traditional memoryless setting; in particular, the distinction between the coding mechanisms for reliability and secrecy becomes apparent in the expressions of the results themselves. In this context, the information-spectrum methods pioneered by Han and Verdú turn out to be convenient mathematical tools, as they allow us to analyze general channels by focusing on the properties of mutual information as a random variable [5], [7], [17]. We note that these tools have already been used to study information-theoretic security beyond memoryless channels and our results provide extensions of [15], [18]–[20].

C. Summary of Results

In this section, we highlight the results presented in this paper, preliminary versions of which have been reported in [21], [22].

- We clarify the relation between information-theoretic security and statistical independence by investigating alternatives to the average mutual information rate $\frac{1}{n}\mathbb{I}(\mathbf{M}; \mathbf{Z}^n)$, which is used as the *de facto* metric in most earlier works. The average mutual information rate is actually a normalized Kullback-Leibler distance between the joint distribution $p_{\mathbf{M}\mathbf{Z}^n}$ and the product distribution $p_{\mathbf{M}}p_{\mathbf{Z}^n}$; the distance between these two distributions can be measured by other means, such as the variational distance or even the cumulative distribution function (CDF) of the random variable $\mathbb{I}(\mathbf{M}; \mathbf{Z}^n)$. By establishing relations among different metrics in Section III, we highlight the importance of choosing a measure of statistical independence that is not only simple enough to be analytically tractable but also strong enough to be cryptographically relevant. In addition, this discussion provides the basis for elegant converse proofs.
- We provide evidence that resolvability may be the fundamental coding mechanism for secure communication by making rigorous the ideas suggested in Example 1 and Example 2. Specifically, we connect secrecy to resolvability to analyze the fundamental limits of Shannon's cipher system (Theorem 1 in Section IV) and of the broadcast channel with confidential messages (Theorem 2 in Section V). In the later case, we show that, at least for a specific wiretap channel, codes

deriving secrecy from resolvability are more powerful than those deriving secrecy from capacity (Proposition 2); we also derive the secrecy capacity region for general broadcast channels with cost constraint and for strong secrecy metrics (Theorem 2 and Theorem 3);

- We further leverage the connection between secrecy and resolvability to revisit various models of secure communication in Section VI. We first provide a simple proof of the strong secrecy capacity of ergodic fading wireless channels with full channel state information [9], [23] (Proposition 3). We then show that achievable rates already known for mixed channels and compound channels [24], [25], can be obtained with conceptually simple proofs, and that these results hold under stronger secrecy metrics than was previously established (Proposition 4 and Proposition 5).
- We exploit the general characterization of secrecy capacity to bound the secret-key capacity of a general discrete source model for secret-key agreement (Theorem 4). This result is obtained by constructing a coding scheme for secret-key agreement from a coding scheme for a wiretap channel. The form of the result, which involves conditional entropy instead of mutual information, suggests that the fundamental mechanism behind secret-key agreement is not resolvability but rather channel intrinsic randomness [26]. Nevertheless, resolvability provides useful insight for secret key agreement. The problem of deriving secrecy from intrinsic randomness is beyond the scope of the present work and will be analyzed in a forthcoming paper.

D. Outline

The remainder of the paper is organized as follows. Section II sets the notation used throughout the paper and briefly reviews the fundamental concepts and results of information-spectrum information theory. Section III introduces and analyzes several secrecy metrics that can be used to measure information-theoretic security. Section IV analyses the fundamental limits of secure communication for Shannon's cipher system. Section V, which forms the core of the paper, proves the impossibility of achieving strong secrecy capacity with random codes deriving secrecy from capacity for some wiretap channels and establishes the secrecy-capacity region of general broadcast channels with confidential messages. Section VI presents applications of the general results to wireless channels, mixed channels and compound channels, and secret-key agreement, which may be of independent interest. Section VII offers some concluding remarks. The technical details of the proofs are organized into a series of lemmas, whose proofs are relegated to the appendices to streamline the presentation.

II. NOTATION AND FOUNDATIONS

To fix notation for the sequel, consider three random variables X , Y , and Z with sample values x , y , and z taking values in alphabets \mathcal{X} , \mathcal{Y} , and \mathcal{Z} , respectively. The joint probability distribution is denoted p_{XYZ} , and the marginal probability distributions are denoted by p_X , p_Y , and p_Z . Unless mentioned otherwise, alphabets are assumed to be abstract alphabets, including countably infinite or continuous alphabets. If the alphabets are finite, then the probability distributions correspond to probability mass functions; if the alphabets are uncountable, then the probability distributions correspond to probability densities, which we assume exist². The *mutual information* between X and Y is the random variable³

$$I(X; Y) \triangleq \log \frac{p_{XY}(X, Y)}{p_X(X) p_Y(Y)}.$$

The average of the mutual information random variable is the usual *average mutual information*, which we denote by $\mathbb{I}(X; Y)$. For discrete random variables, $\mathbb{I}(X; Y)$ has the familiar expression

$$\mathbb{I}(X; Y) \triangleq \mathbb{E}_{XY}[I(X; Y)] = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p_{XY}(x, y) \log \frac{p_{XY}(x, y)}{p_X(x) p_Y(y)}.$$

The conditional mutual information between X and Y given Z and the average conditional mutual information are accordingly defined as

$$I(X; Y|Z) \triangleq \log \frac{p_{XY|Z}(X, Y|Z)}{p_{X|Z}(X|Z) p_{Y|Z}(Y|Z)} \quad \text{and} \quad \mathbb{I}(X; Y|Z) \triangleq \mathbb{E}_{XYZ}[I(X; Y|Z)],$$

respectively. Similarly, the *entropy* and *average entropy* of X are

$$H(X) \triangleq \log \frac{1}{p_X(X)} \quad \text{and} \quad \mathbb{H}(X) \triangleq \mathbb{E}_X[H(X)],$$

and the conditional entropy and average conditional entropy of X given Y are

$$H(X|Y) \triangleq \log \frac{1}{p_{X|Y}(X|Y)} \quad \text{and} \quad \mathbb{H}(X|Y) \triangleq \mathbb{E}_{XY}[H(X|Y)].$$

The binary entropy function is denoted by $\mathbb{H}_b : p \rightarrow -p \log p - (1 - p) \log(1 - p)$. All the traditional relations between average mutual information and average entropy that result from basic properties of joint, marginal, or conditional probability distributions can be shown to hold with probability one for the mutual information and entropy random variables. In particular, the chain rules of mutual information and entropy hold with probability one.

²We note that more general situations can be treated with the approach of Pinsker [27].

³Unless indicated otherwise, logarithms and exponentials in the paper are taken to base two.

In the remainder of the paper, we often measure the similarity of two random variables $X \in \mathcal{X}$ and $X' \in \mathcal{X}$ in terms of the *variational distance* between their distributions, defined as⁴

$$\mathbb{V}(p_X, p_{X'}) \triangleq 2 \sup_{\mathcal{A} \subseteq \mathcal{X}} |\mathbb{P}_X[\mathcal{A}] - \mathbb{P}_{X'}[\mathcal{A}]|.$$

The variational distance is not as convenient to manipulate as the average mutual information, but we provide simple rules for variational distance calculus in Appendix A.

Given two sequences of arbitrary random variables $\{X^n \in \mathcal{X}^n\}_{n \geq 1}$ and $\{Y^n \in \mathcal{Y}^n\}_{n \geq 1}$, characterized by a sequence of joint probability distributions $\{p_{X^n Y^n}\}_{n \geq 1}$, the probability distribution of $\frac{1}{n} I(X^n; Y^n)$ is referred to as the *mutual information rate spectrum*. In addition, the *spectral-inf mutual information rate* is defined as [7]

$$\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y^n) \triangleq \sup \left\{ \beta : \lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{1}{n} I(X^n; Y^n) < \beta \right] = 0 \right\},$$

and the *spectral-sup mutual information rate* is defined as

$$\text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y^n) \triangleq \inf \left\{ \alpha : \lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{1}{n} I(X^n; Y^n) > \alpha \right] = 0 \right\}.$$

For convenience, we recall that these two quantities, which represent the extreme points of the support of the random variable $\frac{1}{n} I(X^n; Y^n)$ in the limit of large n , have an important operational significance for point-to-point communication channels. Given a general channel $(\mathcal{X}, \mathcal{Y}, \{p_{Y^n|X^n}\}_{n \geq 1})$ with input alphabet \mathcal{X} , output alphabet \mathcal{Y} , and transition probabilities $\{p_{Y^n|X^n}\}_{n \geq 1}$, the spectral-inf mutual information rate characterizes the *channel capacity*, defined as the supremum of reliable communication rates over the channel.

Theorem (Verdú-Han [7], [17]). *The channel capacity C of a channel $(\mathcal{X}, \mathcal{Y}, \{p_{Y^n|X^n}\}_{n \geq 1})$ is*

$$C = \max_{\{X^n\}_{n \geq 1}} \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y^n).$$

The spectral-sup mutual information rate characterizes an upper bound for the *channel resolvability*, defined as the infimum rate of uniform randomness required to reproduce any process at the output of the channel with arbitrary precision, measured in terms of variational distance.

Theorem (Han-Verdú [5], [7]). *The channel resolvability S of a channel $(\mathcal{X}, \mathcal{Y}, \{p_{Y^n|X^n}\}_{n \geq 1})$ satisfies*

$$S \leq \max_{\{X^n\}_{n \geq 1}} \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y^n).$$

⁴This general definition of variational distance reduces to $\sum_{x \in \mathcal{X}} |p_X(x) - p_{X'}(x)|$ if \mathcal{X} is countable.

Similarly, given an arbitrary process $\{X^n \in \mathcal{X}^n\}_{n \geq 1}$, the *entropy rate spectrum* is the distribution of the random variable $\frac{1}{n}H(X^n)$, and the spectral-inf entropy rate is defined as

$$\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n}H(X^n) \triangleq \sup \left\{ \beta : \lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{1}{n}H(X^n) < \beta \right] = 0 \right\},$$

while the spectral-sup entropy rate is

$$\text{p-limsup}_{n \rightarrow \infty} \frac{1}{n}H(X^n) \triangleq \inf \left\{ \alpha : \lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{1}{n}H(X^n) > \alpha \right] = 0 \right\}.$$

Again, these two quantities have an operational significance for sources of information. Given an arbitrary source $(\mathcal{X}, \{p_{X^n}\}_{n \geq 1})$ with alphabet \mathcal{X} and symbol sequence probabilities $\{p_{X^n}\}_{n \geq 1}$, the spectral-inf entropy rate represents the *source intrinsic randomness*, that is the maximum rate of uniform randomness that can be extracted from it.

Theorem (Vembu-Verdú [7], [28]). *The source intrinsic randomness S_i of a source $(\mathcal{X}, \{p_{X^n}\}_{n \geq 1})$ is*

$$S_i = \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n}H(X^n).$$

The spectral-sup entropy rate has a dual role and characterizes the *source resolvability*, that is the infimum rate of uniform randomness required to simulate it with arbitrary precision, measured in terms of variational distance.

Theorem (Han-Verdú [5], [7]). *The source resolvability S_r of as source $(\mathcal{X}, \{p_{X^n}\}_{n \geq 1})$ is*

$$S_r = \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n}H(X^n).$$

As we will see, the spectral-sup and spectral-inf mutual information and entropy rates also play fundamental roles in the analysis of secure communications, and many results combine these quantities in various ways.

III. PRELIMINARIES: SECRECY METRICS

Let $n \in \mathbb{N}^*$ and $R > 0$. Let $M \in \llbracket 1, 2^{nR} \rrbracket$ be a random variable with uniform distribution that represents a message in a communication scheme. Assume that an eavesdropper has some knowledge about M represented by another random variable $Z^n \in \mathcal{Z}^n$, characterized by the joint probability distribution p_{MZ^n} . As mentioned in the introduction, message M is information-theoretically secure if it is statistically independent of Z^n ; however, exact statistical independence between M and Z^n is extremely stringent and, for tractability, it is convenient to use a slightly weaker measure of secrecy, by which we only require M and Z^n to be *asymptotically* independent as the parameter n tends to infinity. Note that there

is some leeway in the definition of asymptotic independence because one can choose a particular metric to measure dependence of M and Z^n . For instance, given any distance d for the space of joint probability distributions on $\llbracket 1, 2^{nR} \rrbracket \times \mathcal{Z}^n$, the quantity $d(p_{MZ^n}; p_M p_{Z^n})$ could be used as a metric, and asymptotic statistical independence then amounts to the condition

$$\lim_{n \rightarrow \infty} d(p_{MZ^n}; p_M p_{Z^n}) = 0.$$

In the following, we specify six reasonable choices for secrecy metrics. The first metric measures statistical independence using the Kullback-Leibler divergence:

$$\mathbb{S}_1(p_{MZ^n}, p_M p_{Z^n}) \triangleq \mathbb{D}(p_{MZ^n} \| p_M p_{Z^n}) = \mathbb{I}(M; Z^n).$$

Note that the secrecy condition $\lim_{n \rightarrow \infty} \mathbb{S}_1(p_{MZ^n}, p_M p_{Z^n}) = 0$ is the well-known *strong secrecy* condition [12]. A second metric that is particularly useful is based on the variational distance:

$$\mathbb{S}_2(p_{MZ^n}, p_M p_{Z^n}) \triangleq \mathbb{V}(p_{MZ^n}, p_M p_{Z^n}).$$

For any $\epsilon > 0$, the asymptotic independence of M and Z^n can also be measured in terms of the CDF of $\mathbb{I}(M; Z^n)$:

$$\mathbb{S}_3(p_{MZ^n}, p_M p_{Z^n}) \triangleq \mathbb{P}[\mathbb{I}(M; Z^n) > \epsilon],$$

in which case the secrecy condition

$$\forall \epsilon > 0 \quad \lim_{n \rightarrow \infty} \mathbb{S}_3(p_{MZ^n}, p_M p_{Z^n}) = 0$$

means that the random variable $\mathbb{I}(M; Z^n)$ converges in probability to zero. Finally, we could also use weakened versions of the metrics above by introducing a normalization by a factor of n as

$$\mathbb{S}_4(p_{MZ^n}, p_M p_{Z^n}) \triangleq \frac{1}{n} \mathbb{D}(p_{MZ^n} \| p_M p_{Z^n}) = \frac{1}{n} \mathbb{I}(M; Z^n),$$

$$\mathbb{S}_5(p_{MZ^n}, p_M p_{Z^n}) \triangleq \frac{1}{n} \mathbb{V}(p_{MZ^n}, p_M p_{Z^n}),$$

$$\text{for } \epsilon > 0 \quad \mathbb{S}_6(p_{MZ^n}, p_M p_{Z^n}) \triangleq \mathbb{P}\left[\frac{1}{n} \mathbb{I}(M; Z^n) > \epsilon\right].$$

The secrecy condition $\lim_{n \rightarrow \infty} \mathbb{S}_4(p_{MZ^n}, p_M p_{Z^n}) = 0$ is the *weak secrecy* condition initially introduced by Wyner [1].

Note that the secrecy conditions⁵ $\lim_{n \rightarrow \infty} \mathbb{S}_i(p_{MZ^n}, p_M p_{Z^n}) = 0$ may not be equivalent for all $i \in \llbracket 1, 6 \rrbracket$; by establishing an ordering among the previous metrics, we formalize what it means for a metric to

⁵The limit should be understood for any $\epsilon > 0$ in the case of metrics \mathbb{S}_3 and \mathbb{S}_6 .

be “stronger” than another. Formally, for $i, j \in \llbracket 1, 6 \rrbracket$, we say that \mathbb{S}_i is *stronger* than \mathbb{S}_j (or equivalently that \mathbb{S}_j is *weaker* than \mathbb{S}_i), and we write $\mathbb{S}_i \succeq \mathbb{S}_j$ if and only if

$$\lim_{n \rightarrow \infty} \mathbb{S}_i(p_{\mathbf{M}Z^n}, p_{\mathbf{M}}p_{Z^n}) = 0 \Rightarrow \lim_{n \rightarrow \infty} \mathbb{S}_j(p_{\mathbf{M}Z^n}, p_{\mathbf{M}}p_{Z^n}) = 0.$$

By construction, it is clear that $\mathbb{S}_1 \succeq \mathbb{S}_4$, $\mathbb{S}_2 \succeq \mathbb{S}_5$ and $\mathbb{S}_3 \succeq \mathbb{S}_6$; however, we establish a more precise result.

Proposition 1. *The secrecy metrics \mathbb{S}_i for $i \in \llbracket 1, 6 \rrbracket$ are ordered as follows.*

$$\mathbb{S}_1 \succeq \mathbb{S}_2 \succeq \mathbb{S}_3 \succeq \mathbb{S}_4 \succeq \mathbb{S}_5 \succeq \mathbb{S}_6.$$

Proof: See Appendix B. ■

A direct consequence of Proposition 1 is that any secure communication scheme satisfying the strongest secrecy metric \mathbb{S}_1 automatically satisfies the secrecy metrics \mathbb{S}_i for $i \in \llbracket 2, 6 \rrbracket$. Conversely, any secure communication scheme that does not satisfy the weakest secrecy metric \mathbb{S}_6 cannot satisfy any of the metrics \mathbb{S}_i for $i \in \llbracket 1, 5 \rrbracket$. Therefore, to establish a coding theorem for a secure communication scheme, we can prove achievability for the strongest metric \mathbb{S}_1 and the converse for the weakest metric \mathbb{S}_6 .

Although the ordering in Proposition 1 follows strictly from mathematical properties, the idea that some metrics are stronger than others is also meaningful from a cryptographic perspective. One can construct examples of communication schemes that present obvious security loopholes while still satisfying the weak secrecy metric \mathbb{S}_4 (see for instance the examples in [4], [23], [29]). It is now accepted that information-theoretic results should hold at least under the secrecy metrics⁶ \mathbb{S}_1 or \mathbb{S}_2 .

IV. SHANNON’S CIPHER SYSTEM

As a first illustration of the connection between secrecy and resolvability, we elaborate on Example 1 and revisit Shannon’s cipher system. We consider the model illustrated in Figure 2, in which a message \mathbf{M} uniformly distributed in $\llbracket 1, 2^{nR} \rrbracket$ is to be communicated reliably from a transmitter (Alice) to a legitimate receiver (Bob) in the presence of an eavesdropper (Eve). Alice and Bob have access to a common discrete source of randomness $(\mathcal{K}, \{p_{\mathbf{K}^n}\}_{n \geq 1})$, characterized by an alphabet \mathcal{K} and a sequence of symbol probabilities $\{p_{\mathbf{K}^n}\}_{n \geq 1}$, which is used to encode \mathbf{M} into a codeword $\mathbf{Z} \in \mathcal{Z}$. Bob’s estimate of the message using \mathbf{Z} and the source is denoted by $\hat{\mathbf{M}}$.

⁶The metrics could be further strengthened by imposing an exponential convergence with n ; however, except in the case of exponentially information stable channels [13], such as memoryless channels, we were unable to prove general results with this additional constraint.

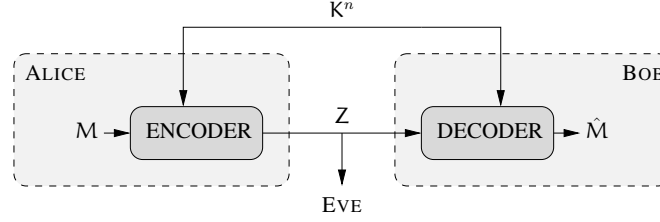


Fig. 2. Shannon's cipher system for a general common source of randomness.

Definition 1. A $(2^{nR}, n)$ code \mathcal{C}_n for Shannon's cipher system consists of

- an encoding function $f_n : \llbracket 1, 2^{nR} \rrbracket \times \mathcal{K}^n \rightarrow \mathcal{Z}$ that encrypts a message into a codeword;
- a decoding function $g_n : \mathcal{Z} \times \mathcal{K}^n \rightarrow \llbracket 1, 2^{nR} \rrbracket$ that decrypts a codeword into a message.

The reliability performance of a code \mathcal{C}_n is measured in terms of the probability of error

$$\mathbb{P}_e(\mathcal{C}_n) \triangleq \mathbb{P}[\hat{M} \neq M | \mathcal{C}_n]$$

while its secrecy performance is measured in terms of the secrecy metric⁷ $\mathbb{S}_i(\mathcal{C}_n) \triangleq \mathbb{S}_i(p_{MZ}, p_{MPZ})$.

Definition 2. A rate R is achievable for secrecy metric \mathbb{S}_i with $i \in \llbracket 1, 6 \rrbracket$ if there exists a sequence of $(2^{nR}, n)$ codes $\{\mathcal{C}_n\}_{n \geq 1}$ such that

$$\lim_{n \rightarrow \infty} \mathbb{P}_e(\mathcal{C}_n) = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \mathbb{S}_i(\mathcal{C}_n) = 0.$$

The secrecy capacity $C_s^{(i)}$ for secrecy metric \mathbb{S}_i is

$$C_s^{(i)} \triangleq \sup\{R : R \text{ is achievable for secrecy metric } \mathbb{S}_i\}.$$

Theorem 1. The secrecy capacity for secrecy metric \mathbb{S}_i with $i \in \llbracket 2, 6 \rrbracket$ is the same and is given by

$$C_s^{(i)} = \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(K^n). \quad (1)$$

If the source $(\mathcal{K}, \{p_{K^n}\}_{n \geq 1})$ is memoryless, then the secrecy capacity is also the same for metric \mathbb{S}_1 .

Proof: We first show that all rates below $\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(K^n)$ are achievable for secrecy metric \mathbb{S}_2 . Let $\epsilon, \gamma > 0$ and $R \triangleq \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(K^n) - \gamma$. Let \mathcal{U}_R be the random variable with uniform distribution on $\llbracket 1, 2^{nR} \rrbracket$. By [28, Lemma 3], there exists an encoding function $f_n : \mathcal{K}^n \rightarrow \llbracket 1, 2^{nR} \rrbracket$ such that $\mathbb{V}(p_{f_n(K^n)}, p_{\mathcal{U}_R}) \leq \epsilon_n$ with $\lim_{n \rightarrow \infty} \epsilon_n = 0$. A message M is then encoded as $Z = f_n(K^n) \oplus M$,

⁷We drop the conditioning on \mathcal{C}_n in probability distributions when this is clear from the context.

where \oplus represents the addition modulo $\lceil 2^{nR} \rceil$. By construction, Bob retrieves M without error since $M = Z \oplus f_n(K^n)$. Then, note that

$$\begin{aligned} \mathbb{S}_2(\mathcal{C}_n) &= \mathbb{V}(p_{MZ}, p_{MpZ}) = \mathbb{E}_M[\mathbb{V}(p_{Z|M}, p_Z)] \\ &\leq \mathbb{E}_M[\mathbb{V}(p_{Z|M}, p_{U_R})] + \mathbb{V}(p_{U_R}, p_Z) \\ &\leq 2\mathbb{E}_M[\mathbb{V}(p_{Z|M}, p_{U_R})] \\ &= 2\mathbb{E}_M[\mathbb{V}(p_{f(K^n)}, p_{U_R})] \\ &\leq 2\epsilon_n, \end{aligned}$$

where the last equality follows from the definition of Z and the independence of $f_n(K^n)$ and M . Therefore, the rate R is achievable and, since γ can be chosen arbitrarily small, we conclude that

$$C_s^{(2)} \geq \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(K^n). \quad (2)$$

If the source $(\mathcal{K}, \{p_{K^n}\}_{n \geq 1})$ is i.i.d., one can easily modify the proof of [28, Lemma 3] to show that, if $R \triangleq \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(K^n) - \gamma$, there exists a function $f_n : \mathcal{K}^n \rightarrow \llbracket 1, 2^{nR} \rrbracket$ and $\alpha_\gamma > 0$, such that $\mathbb{V}(p_{f_n(K^n)}, p_{U_R}) \leq 2^{-\alpha_\gamma n}$. Following the same steps as in the achievability proof above, we then obtain that $\mathbb{S}_2(\mathcal{C}_n) \leq 2 \cdot 2^{-\alpha_\gamma n}$. Then, [13, Lemma 1] shows that there exists $\beta_\gamma > 0$ such that, for n large enough $\mathbb{S}_1(\mathcal{C}_n) \leq 2^{-\beta_\gamma n}$.

We now prove the converse part of the result. Let R be an achievable rate for secrecy metric \mathbb{S}_6 . There exists a sequence of $(2^{nR}, n)$ codes $\{\mathcal{C}_n\}_{n \geq 1}$ such that $\lim_{n \rightarrow \infty} \mathbb{P}_e(\mathcal{C}_n) = 0$ and $\lim_{n \rightarrow \infty} \mathbb{S}_6(\mathcal{C}_n) = 0$. For every $n \in \mathbb{N}^*$, and with probability one, we have

$$\begin{aligned} \frac{1}{n} H(M) &= \frac{1}{n} H(M|Z) + \frac{1}{n} I(M; Z) \\ &= \frac{1}{n} I(M; K^n|Z) + \frac{1}{n} H(M|ZK^n) + \frac{1}{n} I(M; Z) \\ &= \frac{1}{n} H(K^n) - \frac{1}{n} H(K^n|MZ) - \frac{1}{n} I(K^n; Z) + \frac{1}{n} H(M|ZK^n) + \frac{1}{n} I(M; Z). \end{aligned}$$

Since $R = \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(M)$, $\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(K^n|MZ) \geq 0$, and $\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(K^n; Z) \geq 0$, we obtain

$$R \leq \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(K^n) + \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} H(M|ZK^n) + \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(M; Z).$$

By assumption, $\text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(M; Z) = 0$ since $\lim_{n \rightarrow \infty} \mathbb{S}_6(\mathcal{C}_n) = 0$. The Verdú-Han Lemma [7], [17] also guarantees that $\text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} H(M|ZK^n) = 0$; hence, we conclude that

$$C_s^{(6)} \leq \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(K^n). \quad (3)$$

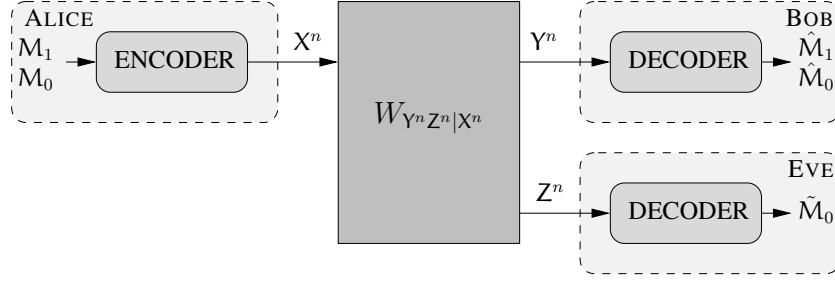


Fig. 3. Broadcast channel with confidential messages.

Combining (2) and (3) with Proposition 1, we conclude that, for each $i \in \llbracket 2, 6 \rrbracket$, $C_s^{(i)} = \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(K^n)$. If the source $(\mathcal{K}, \{p_{K^n}\}_{n \geq 1})$ is i.i.d., then for each $i \in \llbracket 1, 6 \rrbracket$, $C_s^{(i)} = \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(K^n)$. ■

The fact that secrecy capacity is identical for all metrics \mathbb{S}_i with $i \in \llbracket 2, 6 \rrbracket$ suggests that asymptotic statistical independence is indeed a fundamental measure of secrecy.

Note that the coding scheme used in Theorem 1 extracts the *source intrinsic randomness* of $(\mathcal{K}, \{p_{K^n}\}_{n \geq 1})$ to protect the message with a one-time pad. Nevertheless, the message is kept secret from the eavesdropper because the encoder exploits the randomness of the source to control the distribution of the eavesdropper's observation; hence, the coding mechanism for secure communication is closer to channel resolvability, which we confirm in the next section.

V. SECRECY FROM RESOLVABILITY OVER NOISY CHANNELS

We now turn our attention to the problem of secure communication over noisy channels. We consider a broadcast channel with confidential messages $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \{W_{Y^n Z^n | X^n}\}_{n \geq 1})$ characterized by an input alphabet \mathcal{X} , two output alphabets \mathcal{Y} and \mathcal{Z} , and a sequence of transition probabilities $\{W_{Y^n Z^n | X^n}\}_{n \geq 1}$. The channels $(\mathcal{X}, \mathcal{Y}, \{W_{Y^n | X^n}\}_{n \geq 1})$ and $(\mathcal{X}, \mathcal{Z}, \{W_{Z^n | X^n}\}_{n \geq 1})$ obtained from the marginals are called the *main channel* and the *eavesdropper's channel*, respectively. The inputs to the channels are also subject to cost constraint $P \in \mathbb{R}^+$; specifically, there exists a sequence of cost functions $\{c_n\}_{n \geq 1}$ with $c_n : \mathcal{X}^n \rightarrow \mathbb{R}_+$, such that any sequence $x^n \in \mathcal{X}^n$ transmitted through the channel should satisfy $\frac{1}{n} c_n(x^n) \leq P$. Following standard practice, the transmitter is named Alice, the receiver observing output Y is named Bob, and the receiver observing output Z is named Eve. As illustrated in Figure 3, Alice wishes to transmit a common message M_0 to both Bob and Eve and an individual message M_1 for Bob alone, viewing Eve as an eavesdropper for message M_1 . Bob's estimates of the messages are denoted by \hat{M}_0 and \hat{M}_1 while Eve's estimate is denoted by \tilde{M}_0 .

Definition 3. A $(2^{nR_0}, 2^{nR_1}, n)$ wiretap code \mathcal{C}_n consists of

- a common message set $\mathcal{M}_0 = \llbracket 1, 2^{nR_0} \rrbracket$;
- an individual message set $\mathcal{M}_1 = \llbracket 1, 2^{nR_1} \rrbracket$;
- an auxiliary message set $\mathcal{M}'_1 = \llbracket 1, 2^{nR'_1} \rrbracket$, with $R'_1 > 0$,⁸ which is used to randomize the transmission of individual messages;
- a source of local randomness $(\mathcal{R}, p_{\mathcal{R}})$, which can be used to further randomize the encoding process and is only known to Alice;
- an encoding function $f_n : \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}'_1 \times \mathcal{R} \rightarrow \mathcal{X}^n$, such that

$$\forall (m_0, m_1, m'_1, r) \in \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}'_1 \times \mathcal{R} \quad \frac{1}{n} c_n(f_n(m_0, m_1, m'_1, r)) \leq P;$$

- a decoding function $g_n : \mathcal{Y}^n \rightarrow \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}'_1$;
- a decoding function $h_n : \mathcal{Z}^n \rightarrow \mathcal{M}_0$.

The auxiliary message is denoted by M'_1 . All messages M_0, M_1, M'_1 are assumed to be uniformly distributed in their respective sets. The size of the auxiliary message set and the source of local randomness $(\mathcal{R}, p_{\mathcal{R}})$ can be optimized as part of the code design, and the eavesdropper is assumed to know the code \mathcal{C}_n , which includes the statistics $p_{\mathcal{R}}$ of the source of local randomness. In the remainder of the paper, we clearly identify the channel inputs and outputs obtained when using a code \mathcal{C}_n by introducing a bar in the notation of the corresponding random variables. For instance, the random variable representing a codeword chosen in \mathcal{C}_n is denoted \bar{X}^n , those representing the corresponding channel outputs are denoted \bar{Y}^n and \bar{Z}^n , and the joint distribution between $M_0, M_1, \bar{X}^n, \bar{Y}^n, \bar{Z}^n$ is

$$\begin{aligned} \forall (m_0, m_1, x^n, y^n, z^n) \in \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n \\ p_{M_0 M_1 \bar{X}^n \bar{Y}^n \bar{Z}^n}(m_0, m_1, x^n, y^n, z^n) = W_{Y^n Z^n | X^n}(y^n, z^n | x^n) p_{\bar{X}^n | M_0 M_1}(x^n | m_0, m_1) \\ p_{M_0}(m_0) p_{M_1}(m_1). \end{aligned} \quad (4)$$

The reliability of a code \mathcal{C}_n is measured in terms of the average probability of error

$$\mathbb{P}_e(\mathcal{C}_n) \triangleq \mathbb{P} \left[(\hat{M}_0, \hat{M}_1, \hat{M}'_1) \neq (M_0, M_1, M'_1) \text{ or } \tilde{M}_0 \neq M_0 \middle| \mathcal{C}_n \right]$$

while its secrecy is measured in terms of the secrecy metric $\mathbb{S}_i(\mathcal{C}_n) \triangleq \mathbb{S}_i(p_{M_1 \bar{Z}^n}, p_{M_1} p_{\bar{Z}^n})$ for $i \in \llbracket 1, 6 \rrbracket$.

⁸Unlike R_0 and R_1 , which are fixed parameters, we allow R'_1 to vary with n .

Definition 4. A rate pair (R_0, R_1) is achievable for secrecy metric \mathbb{S}_i if there exists a sequence of $(2^{nR_0}, 2^{nR_1}, n)$ codes $\{\mathcal{C}_n\}_{n \geq 1}$ such that

$$\lim_{n \rightarrow \infty} \mathbb{P}_e(\mathcal{C}_n) = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \mathbb{S}_i(\mathcal{C}_n) = 0.$$

The secrecy-capacity region $\mathcal{R}_{BCC}^{(i)}$ for secrecy metric \mathbb{S}_i is

$$\mathcal{R}_{BCC}^{(i)} \triangleq \text{closure}(\{(R_0, R_1) : (R_0, R_1) \text{ is achievable for secrecy metric } \mathbb{S}_i\});$$

the secrecy capacity for secrecy metric \mathbb{S}_i is

$$C_s^{(i)} \triangleq \sup\{R_1 : (0, R_1) \text{ is achievable for secrecy metric } \mathbb{S}_i\}.$$

Note that our definition of a wiretap code explicitly introduces the randomness used in the encoding process. The randomness is split between an auxiliary message with uniform distribution and a source of local randomness and, in addition, we require the auxiliary message to be decoded by the legitimate receiver. Since the source of local randomness can be arbitrarily chosen, our definition incurs no loss of generality; however, this allows us to explicitly define the class of *capacity-based wiretap codes*, which is implicitly used in [1], [2].

Definition 5. A $(2^{nR_0}, 2^{nR_1}, n)$ capacity-based wiretap code \mathcal{C}_n is a $(2^{nR_0}, 2^{nR_1}, n)$ wiretap code such that :

- the auxiliary message rate is $R'_1 = C_e - \epsilon_n$, where C_e is the eavesdropper's channel capacity and $\{\epsilon_n\}_{n \geq 1}$ is such that $\lim_{n \rightarrow \infty} \epsilon_n = 0$ and $\lim_{n \rightarrow \infty} \epsilon_n \sqrt{n} = \infty$;
- there exists an additional decoding function $h'_n : \mathcal{Z}^n \times \mathcal{M}_1 \rightarrow \mathcal{M}'_1$, which allows the eavesdropper to estimate the auxiliary message M'_1 from the observation of \bar{Z}^n and M_1 .

We let \tilde{M}'_1 denote Eve's estimate of M'_1 . The reliability of a capacity-based wiretap code \mathcal{C}_n is then measured in terms of the modified average probability of error

$$\mathbb{P}_e^*(\mathcal{C}_n) \triangleq \mathbb{P}\left[(\hat{M}_0, \hat{M}_1, \hat{M}'_1) \neq (M_0, M_1, M'_1) \text{ or } (\tilde{M}_0, \tilde{M}'_1) \neq (M_0, M'_1) \middle| \mathcal{C}_n\right].$$

Definition 6. A rate pair (R_0, R_1) is achievable for secrecy metric \mathbb{S}_i with capacity-based wiretap codes if there exists a sequence of $(2^{nR_0}, 2^{nR_1}, n)$ capacity-based wiretap codes $\{\mathcal{C}_n\}_{n \geq 1}$ such that

$$\lim_{n \rightarrow \infty} \mathbb{P}_e^*(\mathcal{C}_n) = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \mathbb{S}_i(\mathcal{C}_n) = 0.$$

The constraint $\lim_{n \rightarrow \infty} \mathbb{P}_e^*(\mathcal{C}_n) = 0$ ensures that, given knowledge of \bar{Z}^n and M_1 , the eavesdropper could reliably decode the auxiliary message M'_1 . Nevertheless, since the eavesdropper does not have

access to the message M_1 , this property is solely used to impose structure on the code. The denomination “capacity-based code” is used because the set of codewords associated to a known pair of messages (M_0, M_1) , which can be thought of as a sub-code of rate $R'_1 = C_e - \epsilon_n$, stems from a sequence of capacity-achieving codes for the eavesdropper’s channel. This property, which is formalized in [30, Theorem 1], is implicitly used in most works that show the existence of wiretap codes achieving secrecy rates for metric \mathbb{S}_4 .

Remark 1. Since $\mathbb{P}_e(C_n)$ only depends on the marginals $\{W_{Y^n|X^n}\}_{n \geq 1}$ and $\mathbb{S}_i(C_n)$ only depends on the marginals $\{W_{Z^n|X^n}\}_{n \geq 1}$, the performance of a wiretap code only depends on the marginals; however, this property is lost with capacity-based wiretap codes because $\mathbb{P}_e^*(C_n)$ depends on $\{W_{Y^n Z^n|X^n}\}_{n \geq 1}$.

Remark 2. Csiszár and Körner [2] analyze the fundamental limits of secure communication more precisely by studying the rate-equivocation region (R_0, R_1, R_e) , where $R_e \leq R_1$ represents the equivocation-rate $\frac{1}{n} \mathbb{H}(M_1|\bar{Z}^n)$ of the eavesdropper about the individual message. Unlike the rates R_0 and R_1 , the notion of equivocation depends on the secrecy metric considered; therefore, we restrict ourselves to the special case of full secrecy rates $R_1 = R_e$, for which we can leverage the result of Proposition 1.

In the absence of a common message ($R_0 = 0$), a broadcast channel with confidential messages is concisely called a wiretap channel, and a $(1, 2^{nR_1}, n)$ code is simply denoted as a $(2^{nR_1}, n)$ code.

A. Capacity-Based Wiretap Codes May Not Achieve Strong Secrecy

All the analyzes of wiretap channel models based on capacity-based wiretap codes derive secrecy for metric \mathbb{S}_4 . Additional modifications of the codes based, for instance, on privacy amplification [12], [14] are required to achieve secrecy for stronger metrics. In this section, we show that this may be a fundamental limitation of capacity-based wiretap codes by proving that sequences of random capacity-based wiretap codes that achieve the weak secrecy capacity *cannot* achieve the strong secrecy capacity.

Specifically, we consider a particular wiretap channel, in which the main channel and the eavesdropper’s channel are both binary symmetric channels with respective cross-over probability δ_1 and δ_2 , such that $0 < \delta_1 < \delta_2 < \frac{1}{2}$. We further assume that no cost constraint is imposed ($\forall x^n \in \mathcal{X}^n \ c_n(x^n) = n$ and $P = 1$) and no source of local randomness is available. Information-theoretic proofs using random codes [1] or polar codes [31] show that the absence of source of local randomness incurs no loss of optimality for this channel.

Proposition 2. *Let $\{C_n\}_{n \geq 1}$ be a sequence of $(2^{nR}, n)$ random capacity-based wiretap codes, obtained by generating codeword symbols independently and uniformly at random. Let the rate of the auxiliary message be such that $R' = 1 - \mathbb{H}_b(\delta_2) - \epsilon_n$ and $R + R' = 1 - \mathbb{H}_b(\delta_1) - \epsilon_n$ and assume there is no source of local randomness. Then, there exists $\eta, \alpha > 0$, such that, for n sufficiently large,*

$$\mathbb{P} \left[\mathbb{S}_2(C_n) > \eta, \mathbb{P}_e^*(C_n) \leq 2^{-\frac{1}{2}\epsilon_n n} \text{ and } \mathbb{S}_4(C_n) \leq 2\epsilon_n \right] \geq 1 - 2^{-\alpha n}$$

i.e., with high probability over the random code ensemble, a sequence of capacity-based random codes achieves the weak secrecy capacity but does not achieve the strong secrecy capacity.

Proof: See Appendix C ■

The result of Proposition 2 generalizes to symmetric channels [32] and we conjecture that it also holds for asymmetric channels, as well as non-random codes. Despite its lack of generality, Proposition 2 shows that a random construction with capacity-based wiretap codes is not powerful enough to prove strong secrecy results, which suggests exploiting a more powerful mechanism to ensure secrecy. In the remainder of the paper, we derive secrecy from resolvability and show that such codes do not suffer from the limitations of capacity-based wiretap codes.

Remark 3. *If the main channel is noiseless, Proposition 2 can be strengthened to prove that no capacity-based wiretap code (including non-random codes) achieves secrecy capacity for metrics \mathbb{S}_2 and \mathbb{S}_1 . This fact was noted in [31] for metric \mathbb{S}_1 with a different argument based on results for finite blocklength channel coding [33].*

B. General Broadcast Channels with Confidential Messages and Cost Constraint

In this section, we establish the secrecy-capacity region of a general broadcast channel with confidential messages for secrecy metrics \mathbb{S}_i with $i \in \llbracket 2, 6 \rrbracket$; the alphabets and transition probabilities of the channel $\{W_{Y^n Z^n | X^n}\}_{n \geq 1}$ are arbitrary, so that the model includes continuous channels and channels with memory. Following the conclusions drawn from Proposition 2, we analyze codes that are more powerful than capacity-based wiretap codes and whose secrecy is tied to the notion of resolvability.

Theorem 2. *The secrecy capacity region of a broadcast channel $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \{W_{Y^n Z^n | X^n}\}_{n \geq 1})$ with confidential messages and cost constraint P is the same for secrecy metrics \mathbb{S}_i with $i \in \llbracket 2, 6 \rrbracket$ and is given*

by

$$\mathcal{R}_{\text{BCC}} = \bigcup_{\{\mathbf{U}^n \mathbf{V}^n \mathbf{X}^n\}_{n \geq 1} \in \mathcal{P}} \left\{ (R_0, R_1) : \begin{aligned} & 0 \leq R_0 \leq \min \left(\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{U}^n; \mathbf{Y}^n), \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{U}^n; \mathbf{Z}^n) \right), \\ & 0 \leq R_1 \leq \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{V}^n; \mathbf{Y}^n | \mathbf{U}^n) - \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{V}^n; \mathbf{Z}^n | \mathbf{U}^n) \end{aligned} \right\} \quad (5)$$

where

$$\mathcal{P} \triangleq \left\{ \{\mathbf{U}^n \mathbf{V}^n \mathbf{X}^n\}_{n \geq 1} : \forall n \in \mathbb{N}^* \quad \begin{aligned} & \mathbf{U}^n \rightarrow \mathbf{V}^n \rightarrow \mathbf{X}^n \rightarrow \mathbf{Y}^n \mathbf{Z}^n \text{ forms a Markov chain} \\ & \text{and } \mathbb{P}[\frac{1}{n} c_n(\mathbf{X}^n) \leq P] = 1 \end{aligned} \right\}.$$

Notice that the form of the secrecy capacity region is the natural generalization of that obtained for memoryless channels in [2, Corollary 1]; however, the main channel statistics affect the secure rate R_1 through their “worst realization” ($\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{V}^n; \mathbf{Y}^n | \mathbf{U}^n)$) while the eavesdropper’s channel statistics affect it through their “best realization” ($\text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{V}^n; \mathbf{Z}^n | \mathbf{U}^n)$). Intuitively, as illustrated in Figure 4, this occurs because the worst case for secure communication is when the main channel conveys the smallest information rate to the legitimate receiver while the eavesdropper’s channel leaks the largest information rate to the eavesdropper. It will be apparent in the proof that this asymmetry, which disappears in the case of memoryless channels, happens because the coding mechanisms used to ensure reliability and secrecy are different.

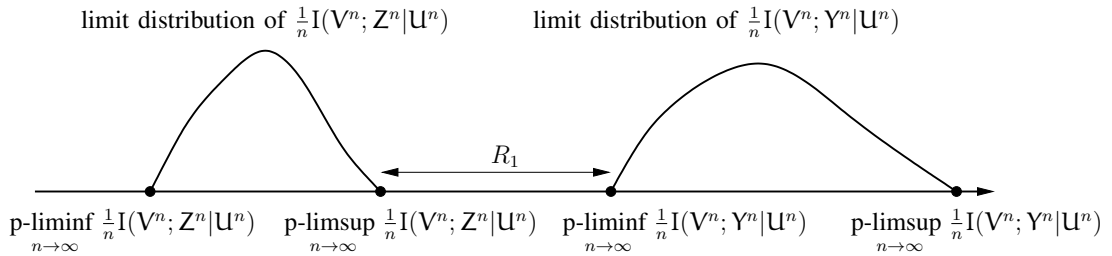


Fig. 4. Illustration of secure rates in Theorem 2.

Without a common message ($R_0 = 0$), we obtain in a similar way the secrecy capacity of a general wiretap channel⁹.

⁹The result in Corollary 1 was already obtained by Hayashi in [15, Lemma 4 and Lemma 5] without cost constraint. Our proof technique is different from that of Hayashi, which is based on a non-asymptotic analysis.

Corollary 1. *The secrecy capacity of a wiretap channel $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \{W_{Y^n Z^n | X^n}\}_{n \geq 1})$ with cost constraint P is identical for secrecy metrics \mathbb{S}_i with $i \in \llbracket 2, 6 \rrbracket$ and is given by*

$$C_s = \max_{\{V^n X^n\}_{n \geq 1} \in \mathcal{P}} \left(\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(V^n; Y^n) - \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(V^n; Z^n) \right), \quad (6)$$

where

$$\mathcal{P} \triangleq \left\{ \{V^n X^n\}_{n \geq 1} : \forall n \in \mathbb{N}^* \quad \begin{array}{l} V^n \rightarrow X^n \rightarrow Y^n Z^n \text{ forms a Markov chain} \\ \text{and } \mathbb{P} \left[\frac{1}{n} c_n(X^n) \leq P \right] = 1 \end{array} \right\}.$$

Remark 4. *The general achievability results of Theorem 2 and Corollary 1 are established for metric \mathbb{S}_2 . We require additional assumptions on the channel statistics to establish secrecy for metric \mathbb{S}_1 , see Remark 5.*

Proof of Theorem 2: We start with the achievability part of the proof, for which we create a codebook by combining superposition coding and binning schemes. Let $n \in \mathbb{N}^*$ and $\epsilon, \gamma, R_0, R_1, R'_1 > 0$. Define $M_0 \triangleq \lceil 2^{nR_0} \rceil$, $M_1 \triangleq \lceil 2^{nR_1} \rceil$ and $M'_1 \triangleq \lceil 2^{nR'_1} \rceil$. Let \mathcal{U} be an arbitrary alphabet and fix a distribution $p_{\mathcal{U}^n}$ on \mathcal{U}^n . Fix a conditional distribution $p_{X^n | \mathcal{U}^n}$ on $\mathcal{X}^n \times \mathcal{U}^n$ such that $\mathbb{P} \left[\frac{1}{n} c_n(X^n) \leq P \right] = 1$. Let $\mathcal{U}^n, X^n, Y^n, Z^n$ be the random variables with joint distribution

$$\forall (u^n, x^n, y^n, z^n) \in \mathcal{U}^n \times \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$$

$$p_{\mathcal{U}^n X^n Y^n Z^n}(u^n, x^n, y^n, z^n) \triangleq W_{Y^n Z^n | X^n}(y^n, z^n | x^n) p_{X^n | \mathcal{U}^n}(x^n | u^n) p_{\mathcal{U}^n}(u^n). \quad (7)$$

- **Code generation:** Randomly generate M_0 sequences $u_k^n \in \mathcal{U}^n$ with $k \in \llbracket 1, M_0 \rrbracket$ according to $p_{\mathcal{U}^n}$. For each $k \in \llbracket 1, M_0 \rrbracket$, generate $M_1 M'_1$ sequence $x_{klm}^n \in \mathcal{X}^n$ with $(l, m) \in \llbracket 1, M_1 \rrbracket \times \llbracket 1, M'_1 \rrbracket$ according to $p_{X^n | \mathcal{U}^n = u_k^n}$. We denote by C_n the random random variable representing the generated code and by \mathcal{C}_n one of its realizations.
- **Encoding:** To transmit a message pair $(k, l) \in \llbracket 1, M_0 \rrbracket \times \llbracket 1, M_1 \rrbracket$, Alice generates an auxiliary message m uniformly at random in $\llbracket 1, M'_1 \rrbracket$ and sends the codeword x_{klm}^n through the channel.
- **Bob's decoding:** Define the sets

$$\begin{aligned} \mathcal{T}_1^n &\triangleq \left\{ (u^n, y^n) \in \mathcal{U}^n \times \mathcal{Y}^n : \frac{1}{n} \log \frac{p_{Y^n | \mathcal{U}^n}(y^n | u^n)}{p_{Y^n}(y^n)} \geq \frac{1}{n} \log M_0 + \gamma \right\} \\ \mathcal{T}_2^n &\triangleq \left\{ (u^n, x^n, y^n) \in \mathcal{U}^n \times \mathcal{X}^n \times \mathcal{Y}^n : \frac{1}{n} \log \frac{p_{Y^n | X^n \mathcal{U}^n}(y^n | x^n, u^n)}{p_{Y^n | \mathcal{U}^n}(y^n | u^n)} \geq \frac{1}{n} \log M_1 M'_1 + \gamma \right\}. \end{aligned}$$

Upon observing y^n , Bob decodes k as the received common message if u_k^n is the unique sequence in \mathcal{C}_n such that $(u_k^n, y^n) \in \mathcal{T}_1^n$; otherwise, a random message is chosen. Similarly, he decodes l as

the received individual message and m as the received auxiliary message if there exists a unique codeword x_{klm}^n such that $(u_k^n, x_{klm}^n, y^n) \in \mathcal{T}_2^n$; otherwise, random messages are chosen.

- **Eve's decoding:** Define the set

$$\mathcal{T}_3^n \triangleq \left\{ (u^n, z^n) \in \mathcal{U}^n \times \mathcal{Z}^n : \frac{1}{n} \log \frac{p_{Z^n|U^n}(z^n|u^n)}{p_{Z^n}(z^n)} \geq \frac{1}{n} \log M_0 + \gamma \right\}$$

Upon observing z^n , Eve decodes k as the received common message if u_k^n is the unique sequence such that $(u_k^n, z^n) \in \mathcal{T}_3^n$; otherwise, a random message is chosen.

The following lemmas, whose proofs are relegated to Appendix D, provide sufficient conditions to guarantee reliability and secrecy.

Lemma 1 (Reliability conditions).

$$\begin{cases} R_0 \leq \min \left(\mathbf{p}\text{-}\liminf_{n \rightarrow \infty} \frac{1}{n} I(U^n; Y^n) - 2\gamma, \mathbf{p}\text{-}\liminf_{n \rightarrow \infty} \frac{1}{n} I(U^n; Z^n) - 2\gamma \right) \\ R_1 + R'_1 \leq \mathbf{p}\text{-}\liminf_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y^n|U^n) - 2\gamma, \end{cases} \Rightarrow \lim_{n \rightarrow \infty} \mathbb{E}[\mathbb{P}_e(C_n)] \leq \epsilon.$$

Lemma 2 (Secrecy from resolvability condition).

$$R'_1 \geq \mathbf{p}\text{-}\limsup_{n \rightarrow \infty} \frac{1}{n} I(X^n; Z^n|U^n) + 2\gamma \Rightarrow \lim_{n \rightarrow \infty} \mathbb{E}[\mathbb{S}_2(C_n)] \leq \epsilon.$$

Combining Lemma 1 and Lemma 2, we obtain

$$\begin{cases} R_0 \leq \min \left(\mathbf{p}\text{-}\liminf_{n \rightarrow \infty} \frac{1}{n} I(U^n; Y^n) - 2\gamma, \mathbf{p}\text{-}\liminf_{n \rightarrow \infty} \frac{1}{n} I(U^n; Z^n) - 2\gamma \right) \\ R_1 \leq \mathbf{p}\text{-}\liminf_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y^n|U^n) - \mathbf{p}\text{-}\limsup_{n \rightarrow \infty} \frac{1}{n} I(X^n; Z^n|U^n) - 4\gamma, \end{cases} \Rightarrow \begin{cases} \lim_{n \rightarrow \infty} \mathbb{E}[\mathbb{P}_e(C_n)] \leq \epsilon \\ \lim_{n \rightarrow \infty} \mathbb{E}[\mathbb{S}_2(C_n)] \leq \epsilon \end{cases}.$$

Using Markov's inequality and the union bound, we can prove there exists at least one sequence of $(2^{nR_0}, 2^{nR_1}, n)$ codes $\{C_n\}_{n \geq 1}$ such that $\lim_{n \rightarrow \infty} \mathbb{P}_e(C_n) \leq 3\epsilon$ and $\lim_{n \rightarrow \infty} \mathbb{S}_2(C_n) \leq 3\epsilon$. Since ϵ and γ can be chosen arbitrarily small, we conclude that

$$\bigcup_{\{U^n X^n\}_{n \geq 1} \in \mathcal{P}} \left\{ (R_0, R_1) : \begin{cases} 0 \leq R_0 \leq \min \left(\mathbf{p}\text{-}\liminf_{n \rightarrow \infty} \frac{1}{n} I(U^n; Y^n), \mathbf{p}\text{-}\liminf_{n \rightarrow \infty} \frac{1}{n} I(U^n; Z^n) \right), \\ 0 \leq R_1 \leq \mathbf{p}\text{-}\liminf_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y^n|U^n) - \mathbf{p}\text{-}\limsup_{n \rightarrow \infty} \frac{1}{n} I(X^n; Z^n|U^n) \end{cases} \right\} \subseteq \mathcal{R}_{\text{BCC}}^{(2)} \quad (8)$$

where

$$\mathcal{P} \triangleq \left\{ \{U^n X^n\}_{n \geq 1} : \forall n \in \mathbb{N}^* \quad \begin{array}{l} U^n \rightarrow X^n \rightarrow Y^n Z^n \text{ forms a Markov chain} \\ \text{and } \mathbb{P}[\frac{1}{n} C_n(X^n) \leq P] = 1 \end{array} \right\}.$$

Finally, note that the source of local randomness (\mathcal{R}, p_R) can be used to prefix an arbitrary channel $(\mathcal{V}, \mathcal{X}, \{p_{X^n|V^n}\}_{n \geq 1})$ to the broadcast channel $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \{W_{Y^n Z^n|X^n}\}_{n \geq 1})$. By applying the proof above

to the concatenated channel $(\mathcal{V}, \mathcal{Y}, \mathcal{Z}, \{p_{Y^n Z^n | V^n}\}_{n \geq 1})$, we conclude that the region given in Theorem 2 is included in the capacity region $\mathcal{R}_{\text{BCC}}^{(2)}$.

We now turn to the converse part of the proof. Consider a sequence of codes $\{\mathcal{C}_n\}_{n \geq 1}$ achieving the rate pair (R_0, R_1) for secrecy metric \mathbb{S}_6 . For $n \in \mathbb{N}^*$, let \bar{U}^n denote the choice of a common message uniformly at random in $[1, 2^{nR_0}]$ and let \bar{W}^n denote the choice of an individual message uniformly at random in $[1, 2^{nR_1}]$. Let \bar{Y}^n and \bar{Z}^n denote the channel outputs corresponding to the transmission of the message pair (\bar{U}^n, \bar{W}^n) . As shown in Appendix E, the following lemmas hold.

Lemma 3.

$$\lim_{n \rightarrow \infty} \mathbb{P}_e(\mathcal{C}_n) = 0 \Rightarrow \begin{cases} R_0 \leq \min \left(\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\bar{U}^n; \bar{Y}^n), \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\bar{U}^n; \bar{Z}^n) \right) \\ R_1 \leq \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\bar{W}^n; \bar{Y}^n | \bar{U}^n). \end{cases}$$

Lemma 4.

$$\begin{cases} \lim_{n \rightarrow \infty} \mathbb{P}_e(\mathcal{C}_n) = 0 \\ \lim_{n \rightarrow \infty} \mathbb{S}_6(\mathcal{C}_n) = 0 \end{cases} \Rightarrow \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(\bar{W}^n; \bar{Z}^n | \bar{U}^n) = 0.$$

Therefore, combining Lemma 3 and Lemma 4, it must hold that

$$\begin{aligned} R_0 &\leq \min \left(\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\bar{U}^n; \bar{Y}^n), \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\bar{U}^n; \bar{Z}^n) \right) \\ R_1 &\leq \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\bar{W}^n; \bar{Y}^n | \bar{U}^n) - \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(\bar{W}^n; \bar{Z}^n | \bar{U}^n). \end{aligned}$$

Note that, by assumption, $\bar{U}^n \bar{W}^n \rightarrow \bar{X}^n \rightarrow \bar{Y}^n \bar{Z}^n$ forms a Markov chain. Define $\bar{V}^n \triangleq (\bar{U}^n, \bar{W}^n)$, which is such that $\bar{U}^n \rightarrow \bar{V}^n \rightarrow \bar{X}^n \rightarrow \bar{Y}^n \bar{Z}^n$ forms a Markov chain. With probability one, we have

$$I(\bar{W}^n; \bar{Y}^n | \bar{U}^n) = I(\bar{V}^n; \bar{Y}^n | \bar{U}^n) \quad \text{and} \quad I(\bar{W}^n; \bar{Z}^n | \bar{U}^n) = I(\bar{V}^n; \bar{Z}^n | \bar{U}^n);$$

therefore, an achievable pair (R_0, R_1) must satisfy

$$\begin{aligned} R_0 &\leq \min \left(\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\bar{U}^n; \bar{Y}^n), \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\bar{U}^n; \bar{Z}^n) \right), \\ \text{and} \quad R_1 &\leq \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\bar{V}^n; \bar{Y}^n | \bar{U}^n) - \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(\bar{V}^n; \bar{Z}^n | \bar{U}^n), \end{aligned}$$

where $\bar{U}^n \rightarrow \bar{V}^n \rightarrow \bar{X}^n \rightarrow \bar{Y}^n \bar{Z}^n$ forms a Markov chain, $p_{\bar{Y}^n \bar{Z}^n | \bar{X}^n} = W_{Y^n Z^n | X^n}$, and $\mathbb{P}[\frac{1}{n} c_n(\bar{X}^n) \leq P] = 1$. Taking the union over all possible processes $\{\bar{U}^n \bar{V}^n \bar{X}^n\}_{n \geq 1}$ gives the desired outer bound for the secrecy capacity region $\mathcal{R}_{\text{BCC}}^{(6)}$.

Since the outer bound for $\mathcal{R}_{\text{BCC}}^{(6)}$ and the inner bound for $\mathcal{R}_{\text{BCC}}^{(2)}$ match, we conclude using Proposition 1 that the secrecy capacity is the same for all metrics $i \in [2, 6]$. ■

A few comments regarding the proof of Theorem 2 are now in order. First, the achievability part of the proof is based on an explicit operational interpretation of secrecy in terms of channel resolvability; in Lemma 2, codes are constructed so that, for a given message M_0 , the probability distribution induced at the eavesdropper's channel output by all messages M_1 is the same. Second, the mechanics of the proof are fundamentally different from the standard approach of Wyner [1] and Csiszár and Körner [2]. The existence of a sequence of codes simultaneously satisfying the reliability and secrecy conditions is obtained by handling the constraints separately, as illustrated by the separate results of Lemma 1 and Lemma 2. This contrasts with the approach of [1], [2], in which the two constraints are handled somewhat simultaneously by using capacity-based wiretap codes. As should be clear from the condition $R'_1 > \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(X^n; Z^n | U^n)$ obtained in Lemma 2, the codes constructed are not capacity-based wiretap codes, for which the condition would read $R'_1 < \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(X^n; Z^n | U^n)$; essentially, channel resolvability allows us to analyze the behavior of codes operating at rates beyond the capacity of the eavesdropper's channel.

Remark 5. *A closer look at the proof of Theorem 2 shows that we could strengthen the secrecy metric and prove that $\mathbb{S}_2(C_n)$ decays exponentially fast with n provided the quantity*

$$\mathbb{P}_{U^n X^n Z^n} \left[\frac{1}{n} I(X^n; Z^n | U^n) > \frac{1}{n} \log M'_1 + \epsilon \right]$$

decays exponentially fast with n for any $\epsilon > 0$.¹⁰ We do not explore this issue further for arbitrary channels but we analyze it more precisely in the next section for memoryless channels.

We conclude by noting that the invariance of the secrecy capacity region with respect to the metrics \mathbb{S}_i for $i \in \llbracket 2, 6 \rrbracket$ suggests that asymptotic statistical independence is indeed a fundamental measure of secrecy because the fundamental limits of secure communication seem to remain unchanged no matter how statistical independence is measured; nevertheless, we emphasize again that practical coding schemes should be designed to provide the strongest level of secrecy.

C. Memoryless Broadcast Channels with Additive Cost Constraint

We now consider memoryless channels (not necessarily discrete) with an additive cost constraint. This is a special case of the general model, in which the transition probabilities factor as

$$\forall (x^n, y^n, z^n) \in \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n \quad W_{Y^n Z^n | X^n}(y^n, z^n | x^n) = \prod_{i=1}^n W_{YZ|X}(y_i, z_i | x_i)$$

¹⁰This property is called exponential information stability in [13].

and the cost constraint satisfies

$$\forall x^n \in \mathcal{X}^n \quad c_n(x^n) = \sum_{i=1}^n c(x_i) \text{ for some cost function } c : \mathcal{X} \rightarrow \mathbb{R}^+.$$

For this special class of channels and constraints and under mild conditions, we can strengthen the results of Section V-B and establish the secrecy capacity region for metric \mathbb{S}_1 .

Theorem 3. *The secrecy-capacity region of a memoryless broadcast channel $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, W_{YZ|X})$ with confidential messages and additive cost constraint P for secrecy metric \mathbb{S}_i with $i \in \llbracket 2, 4 \rrbracket$ is*

$$\mathcal{R}_{\text{BCC}} = \bigcup_{(\mathcal{U}, \mathcal{V}, \mathcal{X}) \in \mathcal{P}} \left\{ (R_0, R_1) : \begin{array}{l} 0 \leq R_0 \leq \min(\mathbb{I}(\mathcal{U}; \mathcal{Y}), \mathbb{I}(\mathcal{U}; \mathcal{Z})) \\ 0 \leq R_1 \leq \mathbb{I}(\mathcal{V}; \mathcal{Y}|\mathcal{U}) - \mathbb{I}(\mathcal{V}; \mathcal{Z}|\mathcal{U}) \end{array} \right\}, \quad (9)$$

where

$$\mathcal{P} \triangleq \{(\mathcal{U}, \mathcal{V}, \mathcal{X}) : \mathcal{U} \rightarrow \mathcal{V} \rightarrow \mathcal{X} \rightarrow \mathcal{Y}\mathcal{Z} \text{ forms a Markov chain and } \mathbb{E}[c(\mathcal{X})] \leq P\}$$

If the rates on the boundary of \mathcal{R}_{BCC} are obtained for some random variables $\mathcal{U}\mathcal{V}\mathcal{X}\mathcal{Y}\mathcal{Z}$ such that the moment generating functions of $\mathbb{I}(\mathcal{V}; \mathcal{Z}|\mathcal{U})$ and $c(\mathcal{X})$ converge unconditionally in a neighborhood of 0 and are differentiable at 0, then \mathcal{R}_{BCC} is also the secrecy-capacity region for \mathbb{S}_1 .

In the absence of a common message ($R_0 = 0$), we obtain in a similar way the following result.

Corollary 2. *The secrecy capacity of a memoryless wiretap channel $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, p_{\mathcal{Y}\mathcal{Z}|\mathcal{X}})$ for secrecy metric \mathbb{S}_i with $i \in \llbracket 2, 4 \rrbracket$ and additive cost constraint P is*

$$C_s = \max_{(\mathcal{V}, \mathcal{X}) \in \mathcal{P}} (\mathbb{I}(\mathcal{V}; \mathcal{Y}) - \mathbb{I}(\mathcal{V}; \mathcal{Z})),$$

where $\mathcal{P} \triangleq \{(\mathcal{V}, \mathcal{X}) : \mathcal{V} \rightarrow \mathcal{X} \rightarrow \mathcal{Y}\mathcal{Z} \text{ forms a Markov chain and } \mathbb{E}[c(\mathcal{X})] \leq P\}$. If the random variables $\mathcal{V}\mathcal{X}\mathcal{Y}\mathcal{Z}$ maximizing C_s are such that the moment generating functions of $\mathbb{I}(\mathcal{V}; \mathcal{Z})$ and $c(\mathcal{X})$ converge unconditionally in a neighborhood of 0 and are differentiable at 0, then C_s is also the secrecy capacity for \mathbb{S}_1 .

Remark 6. For general memoryless channels, the weakest metric for which we show Theorem 3 and Corollary 2 hold is metric \mathbb{S}_4 . As discussed in the proof of Theorem 3, this can be weakened to metric \mathbb{S}_6 for discrete memoryless channels.

Remark 7. The conditions that yield \mathcal{R}_{BCC} and C_s for metric \mathbb{S}_1 are sufficient conditions required to obtain exponential upper bounds when applying Chernov bounds. These conditions are not too restrictive

and are automatically satisfied for discrete memoryless channels and for Gaussian channels with additive power constraint.

Corollary 2 was already obtained for discrete memoryless channels by Csiszár [13] and Maurer and Wolf [12] with different tools. Csiszár's approach uses graph-coloring techniques while Maurer and Wolf's approach exploits privacy amplification with extractors. Theorem 3 for discrete memoryless channels without cost constraint was also obtained independently in [34] using privacy amplification.

Proof of Theorem 3: For discrete memoryless channels, the converse part for secrecy metric \mathbb{S}_6 follows from Theorem 2 without resorting to Fano's inequality. Following [7, Theorem 3.5.2], one can show that

$$\begin{aligned} \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{U}^n; \mathbf{Y}^n) &\leq \liminf_{n \rightarrow \infty} \frac{1}{n} \mathbb{I}(\mathbf{U}^n; \mathbf{Y}^n), \quad \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{U}^n; \mathbf{Z}^n) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \mathbb{I}(\mathbf{U}^n; \mathbf{Z}^n), \\ \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{V}^n; \mathbf{Y}^n | \mathbf{U}^n) &\leq \liminf_{n \rightarrow \infty} \frac{1}{n} \mathbb{I}(\mathbf{V}^n; \mathbf{Y}^n | \mathbf{U}^n), \quad \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{V}^n; \mathbf{Z}^n | \mathbf{U}^n) \geq \limsup_{n \rightarrow \infty} \frac{1}{n} \mathbb{I}(\mathbf{V}^n; \mathbf{Z}^n | \mathbf{U}^n). \end{aligned}$$

Hence, any achievable pair (R_0, R_1) must satisfy

$$\begin{aligned} 0 &\leq R_0 \leq \min \left(\liminf_{n \rightarrow \infty} \frac{1}{n} \mathbb{I}(\mathbf{U}^n; \mathbf{Y}^n), \liminf_{n \rightarrow \infty} \frac{1}{n} \mathbb{I}(\mathbf{U}^n; \mathbf{Z}^n) \right), \\ 0 &\leq R_1 \leq \liminf_{n \rightarrow \infty} \left(\frac{1}{n} \mathbb{I}(\mathbf{V}^n; \mathbf{Y}^n | \mathbf{U}^n) - \frac{1}{n} \mathbb{I}(\mathbf{V}^n; \mathbf{Z}^n | \mathbf{U}^n) \right), \end{aligned}$$

and, for any $\epsilon \geq 0$, we have for n sufficiently large:

$$\begin{aligned} 0 &\leq R_0 \leq \min \left(\frac{1}{n} \mathbb{I}(\mathbf{U}^n; \mathbf{Y}^n), \frac{1}{n} \mathbb{I}(\mathbf{U}^n; \mathbf{Z}^n) \right) + \epsilon, \\ 0 &\leq R_1 \leq \left(\frac{1}{n} \mathbb{I}(\mathbf{V}^n; \mathbf{Y}^n | \mathbf{U}^n) - \frac{1}{n} \mathbb{I}(\mathbf{V}^n; \mathbf{Z}^n | \mathbf{U}^n) \right) + \epsilon. \end{aligned}$$

Setting $\mathbf{M}_0 \triangleq \mathbf{V}^n$ and $\mathbf{M}_1 \triangleq \mathbf{U}^n$, we obtain the same n -letter upper bound as in [2, Section 5]; therefore, the same single-letterization procedure can be applied, which yields the desired result. If the channel alphabets are not discrete, the converse in [2, Section 5] holds for secrecy metric \mathbb{S}_4 .

The achievability of the secrecy-capacity region in Theorem 3 for secrecy metric \mathbb{S}_2 and without cost constraint ($\forall x \in \mathcal{X} \ c(x) = 1$ and $P = 1$) can be directly obtained by substituting appropriate random processes in the general expression of Theorem 2. It suffices to choose i.i.d. processes $\{\mathbf{U}^n \mathbf{V}^n \mathbf{X}^n\}_{n \geq 1}$ such that, for all $n \geq 1$ and for all $(\mathbf{u}, \mathbf{v}, \mathbf{x}) \in \mathcal{U} \times \mathcal{V} \times \mathcal{X}$, $p_{\mathbf{U}_n \mathbf{V}_n \mathbf{X}_n}(\mathbf{u}, \mathbf{v}, \mathbf{x}) = p_{\mathbf{X}|\mathbf{V}}(\mathbf{x}|\mathbf{v}) p_{\mathbf{V}|\mathbf{U}}(\mathbf{v}, \mathbf{u})$; Khintchin's law of large numbers then guarantees that

$$\begin{aligned} \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{U}^n; \mathbf{Y}^n) &= \mathbb{I}(\mathbf{U}; \mathbf{Y}), \quad \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{U}^n; \mathbf{Z}^n) = \mathbb{I}(\mathbf{U}; \mathbf{Z}), \\ \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{V}^n; \mathbf{Y}^n | \mathbf{U}^n) &= \mathbb{I}(\mathbf{V}; \mathbf{Y} | \mathbf{U}), \quad \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{V}^n; \mathbf{Z}^n | \mathbf{U}^n) = \mathbb{I}(\mathbf{V}; \mathbf{Z} | \mathbf{U}), \end{aligned}$$

which yields the desired result; however, additional work is needed to deal with the cost constraint and to obtain secrecy under metric \mathbb{S}_1 . Details are provided in Appendix F. ■

Remark 8. *In the proof of Theorem 3, we actually establish a stronger result than the one stated. If the conditions for the moment generating functions of $I(V; Z|U)$ and $c(X)$ are satisfied, we show that $\mathbb{S}_1(\mathcal{C}_n)$ vanishes exponentially fast with n .*

Remark 9. *Consider a Gaussian wiretap channel with power constraint P , for which $W_{Y|X} \sim \mathcal{N}(0, \sigma_m^2)$ and $W_{Z|X} \sim \mathcal{N}(0, \sigma_e^2)$ with $\sigma_e^2 \geq \sigma_m^2$. Substituting $V = 0$ and $X \sim \mathcal{N}(0, P)$ in Corollary 2, we obtain that all rates R , such that*

$$R < \frac{1}{2} \log \left(1 + \frac{P}{\sigma_m^2} \right) - \frac{1}{2} \log \left(1 + \frac{P}{\sigma_e^2} \right)$$

are achievable secrecy rates for metric \mathbb{S}_1 . Together with the converse proof for metric \mathbb{S}_4 in [8], this establishes the strong secrecy capacity of the Gaussian wiretap channel.

VI. APPLICATIONS

In this section, we illustrate the usefulness of deriving secrecy from resolvability by considering several problems in which the derivation of achievable secrecy rates is tremendously simplified. In particular, results for wireless channels, mixed wiretap channels and compound wiretap channels come almost “for free”. For clarity, we only consider cases in which the common message rate is zero ($R_0 = 0$).

A. Ergodic Wireless Channels with Full Channel State Information

We consider the situation in which Alice and Bob communicate over an ergodic fading wiretap channel and have access to the instantaneous fading gains for both the main channel and the eavesdropper channel. Specifically, at each time $k \geq 1$, the relationships between input and outputs are given by

$$Y_k = H_{m,k} X_k + N_{m,k},$$

$$Z_k = H_{e,k} X_k + N_{e,k},$$

where $\{H_{m,k}\}_{k \geq 1}$, $\{H_{e,k}\}_{k \geq 1}$ are fading gains known to all parties and $\{N_{m,k}\}_{k \geq 1}$, $\{N_{e,k}\}_{k \geq 1}$ are i.i.d. complex Gaussian zero-mean noise processes with respective variance σ_m^2 and σ_e^2 . In addition, channel inputs are subject to the long-term power constraint $\frac{1}{n} \sum_{k=1}^n \mathbb{E}[X_k^2] \leq P$.

Proposition 3. *The secrecy capacity of the ergodic wireless channel with full channel state information for secrecy metric \mathbb{S}_1 is*

$$C_s = \max_{\gamma} \mathbb{E} \left[\log \left(1 + \frac{|H_m|^2 \gamma(H_m, H_e)}{\sigma_m^2} \right) - \log \left(1 + \frac{|H_e|^2 \gamma(H_m, H_e)}{\sigma_e^2} \right) \right], \quad (10)$$

where the maximization is over all power allocation functions $\gamma : \mathbb{C}^2 \rightarrow \mathbb{R}^+$ such that $\mathbb{E}[\gamma(H_m, H_e)] \leq P$.

Proposition 3 states the strong secrecy capacity of wireless channels with full channel state information. This result has already been established in [23] with a completely different approach; deriving secrecy from resolvability and leveraging Corollary 2 provides a much simpler proof, which can be generalized to include the effect of imperfect channel state information [35].

Sketch of proof: We only sketch the achievability part of the proof; the converse for secrecy metric \mathbb{S}_4 is established in [9]. Because the channel gains are instantaneously known to all parties, the ergodic wireless channel can be demultiplexed into a set of independent Gaussian wiretap channels, each characterized by a specific realization (h_m, h_e) of the channel gains and subject to a power constraint $\gamma(h_m, h_e)$. According to Remark 9, the secrecy capacity of each channel for metric \mathbb{S}_1 is

$$\log \left(1 + \frac{|h_m|^2 \gamma(h_m, h_e)}{\sigma_m^2} \right) - \log \left(1 + \frac{|h_e|^2 \gamma(h_m, h_e)}{\sigma_e^2} \right).$$

Hence, using the ergodicity of the channel, we conclude that all the rates R such that

$$0 \leq R < \max_{\gamma} \mathbb{E} \left[\log \left(1 + \frac{|H_m|^2 \gamma(H_m, H_e)}{\sigma_m^2} \right) - \log \left(1 + \frac{|H_e|^2 \gamma(H_m, H_e)}{\sigma_e^2} \right) \right]$$

are achievable for metric \mathbb{S}_1 , where $\gamma : \mathbb{C}^2 \rightarrow \mathbb{R}^+$ satisfies $\mathbb{E}[\gamma(H_m, H_e)] \leq P$. ■

B. Mixed and Compound Channels

As another application of deriving secrecy from resolvability, we analyze the case of mixed and compound wiretap channels. These models have practical relevance since they allow one to analyze situations in which the channel is imperfectly known to the transmitter, either because the channel estimation mechanism is imperfect or because the channel is partially controlled by the eavesdropper.

Let $k \in \mathbb{N}^*$ and let $\{\alpha_k\}_{k \in [1, K]} \in (\mathbb{R}_+^*)^K$ be such that $\sum_{k=1}^K \alpha_k = 1$. Consider K wiretap channels $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \{W_{Y_k^n Z_k^n | X^n}\}_{n=1}^\infty)$ for $k \in [1, K]$. The *mixed wiretap channel* is the channel $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, W_{Y^n Z^n | X^n})$ whose transition probabilities satisfy

$$\forall n \in \mathbb{N}^* \forall (x^n, y^n, z^n) \in \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n \quad W_{Y^n Z^n | X^n}(y^n, z^n | x^n) = \sum_{k=1}^K \alpha_k W_{Y_k^n Z_k^n | X^n}(y^n, z^n | x^n).$$

Proposition 4. *The secrecy capacity of a mixed wiretap channel for secrecy metrics \mathbb{S}_i with $i \in \llbracket 2, 6 \rrbracket$ is*

$$\max_{\{V^n, X^n\}_{n \geq 1} \in \mathcal{P}} \left(\min_{k \in \llbracket 1, K \rrbracket} \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(V^n; Y_k^n) - \max_{k \in \llbracket 1, K \rrbracket} \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(V^n; Z_k^n) \right), \quad (11)$$

where

$$\mathcal{P} \triangleq \left\{ \{V^n X^n\}_{n \geq 1} : \forall n \in \mathbb{N}^* \forall k \in \llbracket 1, K \rrbracket \begin{array}{l} V^n \rightarrow X^n \rightarrow Y_k^n Z_k^n \text{ forms a Markov chain} \\ \text{and } \mathbb{P} \left[\frac{1}{n} c_n(X^n) \leq P \right] = 1 \end{array} \right\}.$$

Proof: Using [7, Lemma 1.4.2], we obtain

$$\begin{aligned} \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(V^n; Y^n) &= \min_{k \in \llbracket 1, K \rrbracket} \left(\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(V^n; Y_k^n) \right) \\ \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(V^n; Z^n) &= \max_{k \in \llbracket 1, K \rrbracket} \left(\text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(V^n; Z_k^n) \right). \end{aligned}$$

The result follows by substituting these equalities in Corollary 1. ■

Note that, for all $i \in \llbracket 1, 6 \rrbracket$, we have $\mathbb{S}_i(p_{M\bar{Z}^n}, p_{MP\bar{Z}^n}) \leq \sum_{k=1}^K \alpha_k \mathbb{S}_i(p_{M\bar{Z}_k^n}, p_{MP\bar{Z}_k^n})$. Therefore, a code ensuring secrecy for the mixed wiretap channel may not guarantee secrecy over each individual wiretap channel. If one wants to ensure secrecy over all possible K channels, one must consider a *compound wiretap channel*, in which the transmitter has no knowledge (even statistical knowledge) of which channel in the set is used for transmission; however, to avoid unnecessary mathematical complications, we assume that receivers can estimate channel statistics perfectly and always know from which channel they obtain observations. For every channel $k \in \llbracket 1, K \rrbracket$, the performance of a code \mathcal{C}_n is measured in terms of the average probability of error $\mathbb{P}_e^{(k)}(\mathcal{C}_n)$ and in terms of the secrecy metric $\mathbb{S}_i^{(k)}(\mathcal{C}_n) \triangleq \mathbb{S}_i(p_{M\bar{Z}_k^n}, p_{MP\bar{Z}_k^n})$; the notion of achievable rate is accordingly modified as follows.

Definition 7. *A rate R is achievable over a compound wiretap channel for secrecy metric \mathbb{S}_i if there exists a sequence of $(2^{nR}, n)$ codes $\{\mathcal{C}_n\}_{n \geq 1}$ such that*

$$\forall k \in \llbracket 1, K \rrbracket \quad \lim_{n \rightarrow \infty} \mathbb{P}_e^{(k)}(\mathcal{C}_n) = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \mathbb{S}_i^{(k)}(\mathcal{C}_n) = 0.$$

Unlike the mixed wiretap channel, there is no distribution associated to the choice of the channel in the set; secrecy and reliability must be guaranteed for all channels in the set, not just the “averaged channel”.

Proposition 5. *The secrecy capacity of a compound wiretap channel with cost constraint P is the same for secrecy metrics \mathbb{S}_i with $i \in \llbracket 2, 6 \rrbracket$ and is given by*

$$\max_{\{V^n, X^n\}_{n \geq 1} \in \mathcal{P}} \left(\min_{k \in \llbracket 1, K \rrbracket} \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(V^n; Y_k^n) - \max_{k \in \llbracket 1, K \rrbracket} \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(V^n; Z_k^n) \right), \quad (12)$$

where

$$\mathcal{P} \triangleq \left\{ \{V^n X^n\}_{n \geq 1} : \forall n \in \mathbb{N}^* \forall k \in \llbracket 1, K \rrbracket \quad \begin{array}{l} V^n \rightarrow X^n \rightarrow Y_k^n Z_k^n \text{ forms a Markov chain} \\ \text{and } \mathbb{P}[\frac{1}{n} c_n(X^n) \leq P] = 1 \end{array} \right\}.$$

Proof: We start with the achievability part of the proof, which is similar to that of Theorem 2. Let $n \in \mathbb{N}^*$ and $\epsilon, \gamma, R_1, R'_1 > 0$. Define $M_1 \triangleq \lceil 2^{nR_1} \rceil$ and $M'_1 \triangleq \lceil 2^{nR'_1} \rceil$. Fix a distribution p_{X^n} on \mathcal{X}^n such that $\mathbb{P}[\frac{1}{n} c_n(X^n) \leq P] = 1$. Let $X^n, \{Y_k^n\}_{k \in \llbracket 1, K \rrbracket}, \{Z_k^n\}_{k \in \llbracket 1, K \rrbracket}$ be the random variables with joint distribution

$$\forall k \in \llbracket 1, K \rrbracket \quad \forall (x^n, y^n, z^n) \in \mathcal{X}^n \times \mathcal{Y}_k^n \times \mathcal{Z}_k^n$$

$$p_{X^n Y_k^n Z_k^n}(x^n, y^n, z^n) \triangleq W_{Y_k^n Z_k^n | X^n}(y^n, z^n | x^n) p_{X^n}(x^n).$$

- **Code generation:** Randomly generate $M_1 M'_1$ sequence $x_{lm}^n \in \mathcal{X}^n$ with $(l, m) \in \llbracket 1, M_1 \rrbracket \times \llbracket 1, M'_1 \rrbracket$ according to p_{X^n} . We denote by C_n the random random variable representing the generated code and by \mathcal{C}_n one of its realizations.
- **Encoding:** To transmit a message $l \in \llbracket 1, M_1 \rrbracket$, Alice generates an auxiliary message m uniformly at random in $\llbracket 1, M'_1 \rrbracket$ and transmits the codeword x_{lm}^n through the channel.
- **Bob's decoding for channel $k \in \llbracket 1, K \rrbracket$:** Define the set

$$\mathcal{T}^n \triangleq \left\{ (x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}_k^n : \frac{1}{n} \log \frac{W_{Y_k^n | X^n}(y^n | x^n)}{p_{Y_k^n}(y^n)} \geq \frac{1}{n} \log M_1 M'_1 + \gamma \right\}.$$

Upon observing y_k^n , Bob decodes l as the received individual message and m as the received auxiliary message if there exists a unique codeword x_{lm}^n such that $(x_{lm}^n, y_k^n) \in \mathcal{T}^n$; otherwise, random messages are chosen.

The following lemmas provide sufficient conditions to guarantee reliability and secrecy. Their proofs are similar to those provided in Appendix D and are omitted.

Lemma 5 (Reliability conditions). *For each $k \in \llbracket 1, K \rrbracket$,*

$$R_1 + R'_1 \leq \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y_k^n) - 2\gamma \Rightarrow \lim_{n \rightarrow \infty} \mathbb{E}[\mathbb{P}_e^{(k)}(C_n)] \leq \epsilon.$$

Lemma 6 (Secrecy from resolvability condition). *For each $k \in \llbracket 1, K \rrbracket$,*

$$R'_1 \geq \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(X^n; Z_k^n) + 2\gamma \Rightarrow \lim_{n \rightarrow \infty} \mathbb{E}[\mathbb{S}_2^{(k)}(C_n)] \leq \epsilon.$$

Using Lemma 5 and 6, we obtain

$$R_1 \leq \min_{k \in \llbracket 1, K \rrbracket} \mathbf{p}\text{-}\liminf_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y_k^n) - \max_{k \in \llbracket 1, K \rrbracket} \mathbf{p}\text{-}\liminf_{n \rightarrow \infty} \frac{1}{n} I(X^n; Z_k^n) - 4\gamma$$

$$\Rightarrow \forall k \in \llbracket 1, K \rrbracket \begin{cases} \lim_{n \rightarrow \infty} \mathbb{E}[\mathbb{P}_e^{(k)}(\mathcal{C}_n)] \leq \epsilon \\ \lim_{n \rightarrow \infty} \mathbb{E}[\mathbb{S}_2^{(k)}(\mathcal{C}_n)] \leq \epsilon \end{cases}.$$

Using Markov's inequality and the union bound, we can show there exists at least one sequence of $(2^{nR_1}, n)$ codes $\{\mathcal{C}_n\}_{n \geq 1}$ such that, for all $k \in \llbracket 1, K \rrbracket$, $\lim_{n \rightarrow \infty} \mathbb{P}_e^{(k)}(\mathcal{C}_n) \leq (K+1)\epsilon$ and $\lim_{n \rightarrow \infty} \mathbb{S}_2^{(k)}(\mathcal{C}_n) \leq (K+1)\epsilon$. Since K is fixed and ϵ, γ can be chosen arbitrarily small, we conclude that all rates R such that

$$0 \leq R_1 < \max_{\{X^n\}_{n \geq 1} \in \mathcal{P}} \left(\min_{k \in \llbracket 1, K \rrbracket} \mathbf{p}\text{-}\liminf_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y_k^n) - \max_{k \in \llbracket 1, K \rrbracket} \mathbf{p}\text{-}\limsup_{n \rightarrow \infty} \frac{1}{n} I(X^n; Z_k^n) \right) \quad (13)$$

are achievable, where $\mathcal{P} \triangleq \{\{X_n\}_{n \geq 1} : \mathbb{P}[\frac{1}{n} c_n(X^n) \leq P] = 1\}$. The achievability of the rates below the secrecy capacity $C_s^{(2)}$ in (12) is then obtained by introducing a prefix channel as in the proof of Theorem 2.

We now turn to the converse part of the proof. Consider a sequence of wiretap codes $\{\mathcal{C}_n\}_{n \geq 1}$ achieving rate R_1 for secrecy metric \mathbb{S}_6 . For $n \in \mathbb{N}^*$, let \bar{V}^n denote the choice of a message uniformly at random in $\llbracket 1, 2^{nR_1} \rrbracket$. By definition, for every $n \in \mathbb{N}^*$ and $k \in \llbracket 1, K \rrbracket$, $\bar{V}^n \rightarrow \bar{X}^n \rightarrow \bar{Y}_k^n \bar{Z}_k^n$ forms a Markov chain and $\mathbb{P}[\frac{1}{n} c(\bar{X}^n) \leq P] = 1$. By the Verdú-Han Lemma [17, Theorem 4], we obtain

$$R < \min_{k \in \llbracket 1, K \rrbracket} \mathbf{p}\text{-}\liminf_{n \rightarrow \infty} \frac{1}{n} I(\bar{V}^n; \bar{Y}_k^n). \quad (14)$$

By definition of the secrecy metric \mathbb{S}_6 , we also have

$$\max_{k \in \llbracket 1, K \rrbracket} \mathbf{p}\text{-}\limsup_{n \rightarrow \infty} \frac{1}{n} I(\bar{V}^n; \bar{Z}_k^n) = 0. \quad (15)$$

Combining (14) and (15), and maximizing over all processes $\{\bar{V}^n \bar{X}^n\}$, we obtain the desired result. ■

Although the secrecy capacity of a compound wiretap channel is identical to that of a mixed wiretap channel, note that the coding schemes achieving it may be fundamentally different. As mentioned earlier, a code designed for a mixed wiretap channel may not guarantee secrecy over the individual channels.

Proposition 6. *Given a memoryless compound wiretap channel with additive cost constraint P , all rates R such that*

$$0 \leq R < \max_{(\mathbf{V}\mathbf{X}) \in \mathcal{P}} \left(\min_{k \in \llbracket 1, K \rrbracket} \mathbb{I}(\mathbf{V}; Y_k) - \max_{k \in \llbracket 1, K \rrbracket} \mathbb{I}(\mathbf{V}; Z_k) \right) \quad (16)$$

are achievable for secrecy metrics \mathbb{S}_i with $i \in \llbracket 2, 6 \rrbracket$, where

$$\mathcal{P} \triangleq \left\{ \mathbf{V}\mathbf{X} : \forall k \in \llbracket 1, K \rrbracket \quad \mathbf{V} \rightarrow \mathbf{X} \rightarrow Y_k Z_k \text{ forms a Markov chain and } \mathbb{E}[c(\mathbf{X})] \leq P \right\}.$$

If the random variables maximizing (16) are such that, for all $k \in \llbracket 1, K \rrbracket$, the moment generating functions of $I(V; Y_k)$ and $c(X)$ converge unconditionally in a neighborhood of 0 and are differentiable at 0, then the rates are also achievable for \mathbb{S}_1 .

Proof: The proof of Proposition 6 follows from steps similar to those used in the proof of Proposition 5 and Theorem 3 and is omitted. ■

When applied to memoryless channels without cost constraint, Proposition 6 provides a generalization of [24, Theorem 1] for strong secrecy. Note that deriving secrecy from resolvability circumvents the enhancement argument used in [24], which is required to show achievability using capacity-based wiretap codes. Similarly, when applied to Gaussian compound wiretap channels with power constraint, Proposition 6 allows one to strengthen [25, Theorem 1].

Remark 10. The general result of Proposition 5 holds provided the number of channels K is fixed and independent of the number n of channel uses; nevertheless, in the special case of Proposition 6, for which we establish secrecy for metric \mathbb{S}_1 , we can show that, for each $k \in \llbracket 1, K \rrbracket$, $\mathbb{S}_1^{(k)} \leq (K+1)2^{-\epsilon_k n}$ for some $\epsilon_k > 0$. Therefore, Proposition 6 also holds if the number of compound channels grows exponentially with n as $K = 2^{\beta n}$ with $\beta < \min_{k \in \llbracket 1, K \rrbracket} \epsilon_k$.

C. Secret-Key Agreement from General Sources.

As a last application, we exploit the result of Corollary 1 to analyze the fundamental limits of secret-key generation for a general source model. Specifically, we consider a *discrete* source $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \{p_{X^n Y^n Z^n}\}_{n \geq 1})$ with three components taking values in discrete alphabets. As illustrated in Figure 5, Alice and Bob attempt to distill a secret-key from their correlated observations X^n and Y^n , respectively, by exchanging messages over a public authenticated channel with unlimited capacity.

Definition 8. A $(2^{nR}, n)$ key-distillation strategy \mathcal{S}_n consists of:

- a key alphabet $\mathcal{K} = \llbracket 1, 2^{nR} \rrbracket$;
- an alphabet \mathcal{A} used by Alice to communicate over the public channel;
- an alphabet \mathcal{B} used by Bob to communicate over the public channel;
- a source of local randomness for Alice (\mathcal{R}_X, p_{R_X}) ;
- a source of local randomness for Bob (\mathcal{R}_Y, p_{R_Y}) ;
- an integer $r \in \mathbb{N}^*$ that represents the number of rounds of communication;
- r encoding functions $f_i : \mathcal{X}^n \times \mathcal{B}^{i-1} \times \mathcal{R}_X \rightarrow \mathcal{A}$ for $i \in \llbracket 1, r \rrbracket$;

- r encoding functions $g_i : \mathcal{Y}^n \times \mathcal{A}^{i-1} \times \mathcal{R}_Y \rightarrow \mathcal{B}$ for $i \in \llbracket 1, r \rrbracket$;
- a key-distillation function $\kappa_a : \mathcal{X}^n \times \mathcal{B}^r \times \mathcal{R}_X \rightarrow \mathcal{K}$;
- a key-distillation function $\kappa_b : \mathcal{Y}^n \times \mathcal{A}^r \times \mathcal{R}_Y \rightarrow \mathcal{K}$;

and operates as follows:

- Alice observes n realizations of the source x^n while Bob observes y^n and Eve observes z^n ;
- Alice generates a realization r_x of her source of local randomness while Bob generates r_y from his;
- in round $i \in \llbracket 1, r \rrbracket$, Alice transmits $a_i = f_i(x^n, b^{i-1}, r_x)$ while Bob transmits $b_i = g_i(y^n, a^{i-1}, r_y)$;
- after round r , Alice computes a key $k = \kappa_a(x^n, b^r, r_x)$ while Bob computes a key $\hat{k} = \kappa_b(y^n, a^r, r_y)$.

The random variables corresponding to r_x and r_y are denoted by R_X and R_Y , respectively; for $i \in \llbracket 1, r \rrbracket$, those corresponding to messages a_i and b_i are denoted by A_i and B_i . The performance of a secret-key distillation strategy \mathcal{S}_n is measured in terms of the average probability of error $\mathbb{P}_e(\mathcal{S}_n) \triangleq \mathbb{P}[\mathbf{K} \neq \hat{\mathbf{K}} | \mathcal{S}_n]$, the secrecy of the key $\mathbb{S}_i(\mathcal{S}_n) \triangleq \mathbb{S}_i(p_{\mathbf{K}Z^n A^r B^r}, p_{\mathbf{K}} p_{Z^n A^r B^r})$ for $i \in \llbracket 1, 6 \rrbracket$, and the uniformity of the key $\mathbb{U}(\mathcal{S}_n) \triangleq 2^{nR} - \mathbb{H}(\mathbf{K})$.

Definition 9. A key rate R is achievable for secrecy metric \mathbb{S}_i if there exists a sequence $\{\mathcal{S}_n\}_{n \geq 1}$ of $(2^{nR}, n)$ key-distillation strategies such that

$$\lim_{n \rightarrow \infty} \mathbb{P}_e(\mathcal{S}_n) = 0, \quad \lim_{n \rightarrow \infty} \mathbb{S}_i(\mathcal{S}_n) = 0, \quad \lim_{n \rightarrow \infty} \mathbb{U}(\mathcal{S}_n) = 0.$$

The secret-key capacity for metric \mathbb{S}_i is defined as

$$C_{SK}^{(i)} \triangleq \sup\{R : R \text{ is an achievable key rate for metric } \mathbb{S}_i\}.$$

Theorem 4. The secret-key capacity of a discrete source $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \{p_{X^n Y^n Z^n}\}_{n \geq 1})$ for secrecy metrics

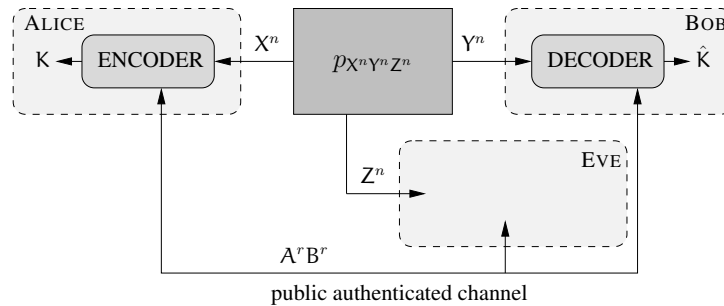


Fig. 5. Secret-key agreement from general source.

\mathbb{S}_i with $i \in \llbracket 2, 6 \rrbracket$ satisfies

$$\begin{aligned} \max \left(\begin{array}{l} \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(X^n | Z^n) - \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} H(X^n | Y^n) \\ \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(Y^n | Z^n) - \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} H(Y^n | X^n) \end{array} \right) &\leq C_{SK}^{(i)} \\ &\leq \min \left(\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y^n), \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y^n | Z^n) \right). \quad (17) \end{aligned}$$

If the discrete source is i.i.d., the above inequalities hold for secrecy metric \mathbb{S}_1 , as already known from [12], [13].

Corollary 3. *The secret-key capacity of an i.i.d discrete source $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, p_{XYZ})$ for secrecy metric \mathbb{S}_1 satisfies*

$$\max(\mathbb{I}(X; Y) - \mathbb{I}(X; Z), \mathbb{I}(X; Y) - \mathbb{I}(Y; Z)) \leq C_{SK}^{(1)} \leq \min(\mathbb{I}(X; Y), \mathbb{I}(X, Y | Z)).$$

Proof of Theorem 4 and Corollary 3: The achievability part of Theorem 4 is based on the construction of a conceptual wiretap channel as in [10]. Assume that Alice, Bob and Eve observe n realizations X^n , Y^n and Z^n of the source, respectively. Consider an arbitrary process $\{U_j\}_{j \geq 1}$ such that $U_j \in \mathcal{X}$ for $j \in \llbracket 1, n \rrbracket$. Assume that Alice forms the signal $U^n \oplus X^n$ on the public channel, in which \oplus denotes the symbol-wise modulo- \mathcal{X} addition. This operation creates a conceptual wiretap channel with input U^n , in which Bob observes the outputs Y^n and $U^n \oplus X^n$ while Eve observes the outputs Z^n and $U^n \oplus X^n$. From Corollary 1, the secrecy capacity of this conceptual channel for secrecy metrics \mathbb{S}_i with $i \in \llbracket 2, 6 \rrbracket$ is at least

$$\max_{U^n} \left(\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(U^n; Y^n, U^n \oplus X^n) - \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(U^n; Z^n, U^n \oplus X^n) \right).$$

In particular, we can choose for $\{U_j\}_{j \geq 1}$ an i.i.d. process such that, for all $j \in \mathbb{N}^*$, U_j is independent of $X^n Y^n Z^n$ and uniformly distributed in \mathcal{X} . Then, with probability one,

$$\begin{aligned} I(U^n; Y^n, U^n \oplus X^n) &= \log \frac{p_{U^n \oplus X^n, Y^n | U^n}(U^n \oplus X^n, Y^n | U^n)}{p_{U^n \oplus X^n, Y^n}(U^n \oplus X^n, Y^n)} \\ &= \log \frac{p_{X^n | Y^n U^n}(X^n | Y^n U^n) p_{Y^n | U^n}(Y^n | U^n)}{p_{U^n \oplus X^n | Y^n}(U^n \oplus X^n | Y^n) p_{Y^n}(Y^n)} \\ &= \log p_{X^n | Y^n}(X^n | Y^n) - \log \frac{1}{|\mathcal{X}|^n}, \end{aligned}$$

where the last inequality follows from $p_{Y^n | U^n}(Y^n | U^n) = p_{Y^n}(Y^n)$, $p_{X^n | Y^n U^n}(X^n | Y^n U^n) = p_{X^n | Y^n}(X^n | Y^n)$ since U^n is independent of $X^n Y^n$ and $p_{U^n \oplus X^n | Y^n}(U^n \oplus X^n | Y^n) = \frac{1}{|\mathcal{X}|^n}$ by the crypto lemma [6].

Therefore,

$$\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{U}^n; \mathbf{Y}^n, \mathbf{U}^n \oplus \mathbf{X}^n) = \log |\mathcal{X}| - \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} H(\mathbf{X}^n | \mathbf{Y}^n). \quad (18)$$

Similarly, one obtains

$$\text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{U}^n; \mathbf{Z}^n, \mathbf{U}^n \oplus \mathbf{X}^n) = \log |\mathcal{X}| - \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(\mathbf{X}^n | \mathbf{Z}^n). \quad (19)$$

Combining (18) and (19), we conclude that any rate R such that

$$R < \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(\mathbf{X}^n | \mathbf{Z}^n) - \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} H(\mathbf{X}^n | \mathbf{Y}^n)$$

is an achievable rate for the conceptual wiretap channel. Since this channel allows one to transmit uniformly distributed messages, R is also an achievable secret-key rate for the source model. The second lower bound is obtained by reversing the role of \mathbf{X}^n and \mathbf{Y}^n in the steps above. For i.i.d. discrete sources, a similar proof based on Corollary 2 in place of Corollary 1 shows that the result holds for metric \mathbb{S}_1 as well.

The proof of the converse is relegated to Appendix G. ■

Remark 11. *Theorem 4 is easily adapted to a “channel model”, in which \mathbf{X}^n is controlled by Alice and broadcasted to Bob and Eve through a channel $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \{p_{\mathbf{Y}^n \mathbf{Z}^n | \mathbf{X}^n}\}_{n \geq 1})$. In this case, the bounds in Theorem 4 include a maximization over all possible distributions $p_{\mathbf{X}^n}$.*

Notice that the general form of the achievable key rates obtained in Theorem 4 involves conditional entropy; except in some special cases, such as i.i.d. sources, this is fundamentally different from the general form of achievable secrecy rates for wiretap channels obtained in Corollary 1, which involves mutual information. In particular, if $\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(\mathbf{X}^n) = \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} H(\mathbf{X}^n)$, then,

$$\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(\mathbf{X}^n | \mathbf{Z}^n) - \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} H(\mathbf{X}^n | \mathbf{Y}^n) \geq \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{X}^n; \mathbf{Y}^n) - \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{X}^n; \mathbf{Z}^n).$$

This distinction suggests that the fundamental mechanism for secret-key distillation, which one would have to exploit to design secret-key distillation strategies without relying on the existence of wiretap codes, is not linked to resolvability; indeed, it has been argued that the mechanism behind secret-key distillation is the *channel intrinsic randomness* [26]. In that respect, despite its generality, the proof of Theorem 4 does not provide much insight into the design of practical secret-key distillation strategies.

VII. CONCLUSION

We have analyzed several models of secure communication over noisy channels by exploiting the idea that the fundamental coding mechanism to ensure secrecy is related to resolvability. This approach has allowed us to establish results for generic channels and for stronger secrecy metrics than the usual average mutual information rate between messages and eavesdropper's observations.

From a technical point of view, deriving secrecy from resolvability provides a conceptually simple approach to analyze the secure achievable rates of many models. Although we have limited examples of applications to mixed, compound, and wireless channels, the connection between secrecy and resolvability is useful in many other settings. Examples of secure communication models for which deriving secrecy from resolvability simplifies the analysis include queuing channels [36], wireless channels with imperfect state information [35], runlength-limited channels [37], and two-way wiretap channels [38].

From a practical perspective, we believe that the connection between strong secrecy and resolvability is fundamental. We have provided evidence of this connection by proving that, for binary symmetric wiretap channels, sequences of random capacity-based wiretap codes, which are implicitly used in [1], [2], cannot achieve the strong secrecy capacity. Although this result has a limited scope, it is consistent with practical code constructions achieving strong secrecy rates [31], [39] and other approaches based on privacy amplification [12], [34].

Our results can be extended in several directions. For instance, the coding mechanisms for secrecy presented in Section IV for Shannon's cipher system and in Section V for wiretap channels can be combined without much difficulty using a coding scheme similar to that proposed in [40]. One could also further investigate the nature of the coding mechanisms for secrecy in secret-key agreement models. Some results along these lines are already available, for instance in [13], [26], [41].

APPENDIX A

TECHNICAL LEMMAS

Lemma 7 (Chernov bound). *Let X be a real valued random variable with moment generating function $\phi_X : \mathbb{R} \rightarrow \mathbb{R} : s \mapsto \mathbb{E}[e^{sX}]$. Let $\{X_i\}_{i=1}^n$ be i.i.d. with distribution p_X . If ϕ_X converges unconditionally in a neighborhood of 0 and is differentiable at 0 then,*

$$\forall \epsilon > 0 \exists \alpha_\epsilon > 0 \text{ such that } \mathbb{P} \left[\frac{1}{n} \sum_{i=1}^n X_i > \mathbb{E}[X] + \epsilon \right] \leq 2^{-\alpha_\epsilon n}.$$

Lemma 8 (Basic properties of variational distance). *Let X_1 , X_2 , and X_3 be random variables defined on the same alphabet \mathcal{X} . Then,*

$$\begin{aligned} \mathbb{V}(p_{X_1}; p_{X_3}) &\leq \mathbb{V}(p_{X_1}; p_{X_2}) + \mathbb{V}(p_{X_2}; p_{X_3}), \\ \text{and } \mathbb{V}(p_{X_1}; p_{X_2}) &\leq \mathbb{V}(p_{X_1} p_{X_3}; p_{X_2} p_{X_3}) = \mathbb{E}_{X_3} [\mathbb{V}(p_{X_1}, p_{X_2|X_3})]. \end{aligned}$$

Proof: The statements are immediate consequences of the definition of variational distance. ■

Lemma 9 (Data-processing inequality for variational distance). *Let X_1 and X_2 be random variables defined on the same alphabet \mathcal{X} . Let $W_{Z|X}$ be transition probabilities from \mathcal{X} to \mathcal{Z} and define the random variables Z_1 and Z_2 such that*

$$\forall (z, x) \in \mathcal{Z} \times \mathcal{X} \quad p_{Z_1 X_1}(z, x) = W_{Z|X}(z|x) p_{X_1}(x) \text{ and } p_{Z_2 X_2}(z, x) = W_{Z|X}(z|x) p_{X_2}(x).$$

Then, $\mathbb{V}(p_{Z_1}, p_{Z_2}) \leq \mathbb{V}(p_{X_1}, p_{X_2})$.

Proof: Note that

$$\begin{aligned} \mathbb{V}(p_{Z_1}, p_{Z_2}) &= \sum_{z \in \mathcal{Z}} |p_{Z_1}(z) - p_{Z_2}(z)| = \sum_{z \in \mathcal{Z}} \left| \sum_{x \in \mathcal{X}} p_{Z|X}(z|x) p_{X_1}(x) - \sum_{x \in \mathcal{X}} p_{Z|X}(z|x) p_{X_2}(x) \right| \\ &\leq \sum_{z \in \mathcal{Z}} \sum_{x \in \mathcal{X}} p_{Z|X}(z|x) |p_{X_1}(x) - p_{X_2}(x)| \\ &= \mathbb{V}(p_{X_1}, p_{X_2}) \end{aligned}$$
■

APPENDIX B

PROOF OF PROPOSITION 1

The fact that $\mathbb{S}_1 \succeq \mathbb{S}_2$ and $\mathbb{S}_4 \succeq \mathbb{S}_5$ follows directly from Pinsker's inequality [27, Corollary p.16]. Similarly, the fact that $\mathbb{S}_2 \succeq \mathbb{S}_3$ and $\mathbb{S}_5 \succeq \mathbb{S}_6$ follows from [27, Corollary p.18]; hence, we only need to prove that $\mathbb{S}_3 \succeq \mathbb{S}_4$.

Let $\epsilon, \gamma > 0$. Assume that $\lim_{n \rightarrow \infty} \mathbb{S}_3(p_{M Z^n}, p_M p_{Z^n}) = 0$, so that $\lim_{n \rightarrow \infty} \mathbb{P}[\mathbb{I}(M; Z^n) > \epsilon] = 0$.

Note that metric $\mathbb{S}_4(p_{\mathbf{M}Z^n}, p_{\mathbf{M}}p_{Z^n})$ can be written as

$$\begin{aligned}
\mathbb{S}_4(p_{\mathbf{M}Z^n}, p_{\mathbf{M}}p_{Z^n}) &= \frac{\mathbb{I}(\mathbf{M}; Z^n)}{n} \\
&= \mathbb{E} \left[\frac{\mathbb{I}(\mathbf{M}; Z^n)}{n} \right], \\
&= \mathbb{E} \left[\frac{\mathbb{I}(\mathbf{M}; Z^n)}{n} \mathbb{1}_{\{\mathbb{I}(\mathbf{M}; Z^n) \leq -\epsilon\}} \right] + \mathbb{E} \left[\frac{\mathbb{I}(\mathbf{M}; Z^n)}{n} \mathbb{1}_{\{-\epsilon < \mathbb{I}(\mathbf{M}; Z^n) \leq \epsilon\}} \right] \\
&\quad + \mathbb{E} \left[\frac{\mathbb{I}(\mathbf{M}; Z^n)}{n} \mathbb{1}_{\{\epsilon < \mathbb{I}(\mathbf{M}; Z^n) \leq n(R + \gamma)\}} \right] + \mathbb{E} \left[\frac{\mathbb{I}(\mathbf{M}; Z^n)}{n} \mathbb{1}_{\{\mathbb{I}(\mathbf{M}; Z^n) > n(R + \gamma)\}} \right].
\end{aligned}$$

Clearly, it holds that

$$\begin{aligned}
\mathbb{E} \left[\frac{\mathbb{I}(\mathbf{M}; Z^n)}{n} \mathbb{1}_{\{\mathbb{I}(\mathbf{M}; Z^n) \leq -\epsilon\}} \right] &< 0, \\
\mathbb{E} \left[\frac{\mathbb{I}(\mathbf{M}; Z^n)}{n} \mathbb{1}_{\{-\epsilon < \mathbb{I}(\mathbf{M}; Z^n) \leq \epsilon\}} \right] &\leq \mathbb{E} \left[\frac{|\mathbb{I}(\mathbf{M}; Z^n)|}{n} \mathbb{1}_{\{|\mathbb{I}(\mathbf{M}; Z^n)| \leq \epsilon\}} \right] \leq \frac{\epsilon}{n},
\end{aligned}$$

and

$$\mathbb{E} \left[\frac{\mathbb{I}(\mathbf{M}; Z^n)}{n} \mathbb{1}_{\{\epsilon < \mathbb{I}(\mathbf{M}; Z^n) \leq n(R + \gamma)\}} \right] \leq (R + \gamma) \mathbb{P}[\mathbb{I}(\mathbf{M}; Z^n) > \epsilon].$$

Following [7, p. 223], we can prove that

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\frac{\mathbb{I}(\mathbf{M}; Z^n)}{n} \mathbb{1}_{\{\mathbb{I}(\mathbf{M}; Z^n) > n(R + \gamma)\}} \right] = 0.$$

Therefore, $\lim_{n \rightarrow \infty} \mathbb{S}_4(p_{\mathbf{M}Z^n}, p_{\mathbf{M}}p_{Z^n}) = 0$ and $\mathbb{S}_3 \succeq \mathbb{S}_4$.

APPENDIX C

PROOF OF PROPOSITION 2

Let C_n be the random variable that denotes a randomly generated code. The proof of the proposition relies on the following two lemmas.

Lemma 10. *There exists $\alpha_1 > 0$, such that, for n sufficiently large,*

$$\mathbb{P} \left[\mathbb{P}_e^*(C_n) \leq 2^{-\frac{1}{2}\epsilon_n n} \text{ and } \mathbb{S}_4(C_n) \leq 2\epsilon_n \right] \geq 1 - 2^{-\alpha_1 n}.$$

Proof: The existence of $\alpha_1 > 0$ such that $\mathbb{P}\left[\mathbb{P}_e^*(\mathcal{C}_n) \leq 2^{-\frac{1}{2}\epsilon_n n}\right] \geq 1 - 2^{-\alpha_1 n}$ follows from a standard random coding argument. Consider a code \mathcal{C}_n such that $\mathbb{P}_e^*(\mathcal{C}_n) \leq 2^{-\frac{1}{2}\epsilon_n n}$. Then, for n large enough,

$$\begin{aligned} \mathbb{S}_4(\mathcal{C}_n) &= \frac{1}{n} \mathbb{I}(\mathbf{M}; \bar{\mathbf{Z}}^n) = \frac{1}{n} \mathbb{I}(\mathbf{M} \bar{\mathbf{X}}^n; \bar{\mathbf{Z}}^n) - \frac{1}{n} \mathbb{I}(\bar{\mathbf{X}}^n; \bar{\mathbf{Z}}^n | \mathbf{M}) \\ &= \frac{1}{n} \mathbb{I}(\bar{\mathbf{X}}^n; \bar{\mathbf{Z}}^n) - \frac{1}{n} \mathbb{H}(\bar{\mathbf{X}}^n; | \mathbf{M}) + \frac{1}{n} \mathbb{H}(\bar{\mathbf{X}}^n | \mathbf{M} \bar{\mathbf{Z}}^n) \\ &\stackrel{(a)}{\leq} C_e - (C_e - \epsilon_n) + R' \mathbb{P}_e^*(\mathcal{C}_n) \\ &\leq 2\epsilon_n. \end{aligned}$$

where (a) follows from Fano's inequality. ■

Let \mathbf{Z}^n be the random variable with uniform distribution on \mathcal{Z}^n , i.e. for every $z^n \in \mathcal{Z}^n$, $p_{\mathbf{Z}^n}(z^n) = \frac{1}{2^n}$.

Lemma 11. *There exists $\beta, \alpha_2 > 0$, such that, for n large enough*

$$\mathbb{P}\left[\mathbb{V}(p_{\bar{\mathbf{Z}}^n}, p_{\mathbf{Z}^n}) \leq 2^{-\beta n}\right] \geq 1 - 2^{-\alpha_2 n}.$$

Proof: This result follows, for instance, from [42, Lemma 19]. ■

For $n \in \mathbb{N}^*$, let \mathcal{C}_n denote a randomly generated code such that

$$\mathbb{P}_e^*(\mathcal{C}_n) \leq 2^{-\frac{1}{2}\epsilon_n n}, \mathbb{S}_4(\mathcal{C}_n) \leq 2\epsilon_n, \text{ and } \mathbb{V}(p_{\bar{\mathbf{Z}}^n}, p_{\mathbf{Z}^n}) \leq 2^{-\beta n}. \quad (20)$$

For n large enough, Lemma 10 and Lemma 11 guarantee that this occurs with probability at least $1 - 2^{-\alpha_3 n}$ with $\alpha_3 = \frac{1}{2} \min(\alpha_1, \alpha_2)$. With a slight abuse of notation, we also let $\mathcal{C}_n \subset \mathcal{X}^n$ denote the codebook and let $f_n^{-1} : \mathcal{C}_n \rightarrow \mathcal{M}_1$ be the restriction to \mathcal{M}_1 of the inverse mapping of f_n . Define new functions ϕ_n and ψ_n as

$$\begin{aligned} \phi_n : \mathcal{C}_n &\rightarrow \mathcal{M}_1 & \text{and} & \quad \psi_n : \mathcal{Z}^n \times \mathcal{M}_1 \rightarrow \mathcal{C}_n \\ x^n &\mapsto f_n^{-1}(x^n) & (z^n, m) &\mapsto f_n(m, h_n(z^n, m)) \end{aligned}$$

These functions define a source code for the compression of the source $\bar{\mathbf{X}}^n \in \mathcal{C}_n$ (the choice of codewords uniformly at random in the code) with $\bar{\mathbf{Z}}^n$ as correlated side information at the receiver, whose probability of decoding error is $\mathbb{P}_e^*(\mathcal{C}_n)$. We now leverage the results obtained by Hayashi [43] and generalized by Watanabe *et al.* [41] that establish a tradeoff between probability and error and resolvability for source coding of arbitrary sources. Combining [41, Theorem 6] and the proof of [41, Theorem 7], we obtain, for any $b > 0$,

$$\forall n \in \mathbb{N}^* \quad \mathbb{P}_e^*(\mathcal{C}_n) + \mathbb{S}_2(\mathcal{C}_n) \geq 1 - \left(2^{-b\sqrt{n}+1} + \mathbb{P}_{\bar{\mathbf{X}}^n \bar{\mathbf{Z}}^n}[\mathcal{A}_0]\right), \quad (21)$$

with

$$\mathcal{A}_0 \triangleq \left\{ (x^n, z^n) \in \mathcal{C}_n \times \mathcal{Z}^n : \frac{2^{-b\sqrt{n}}}{|\mathcal{M}_1|} < p_{\bar{X}^n|\bar{Z}^n}(x^n|z^n) \leq \frac{2^{b\sqrt{n}}}{|\mathcal{M}_1|} \right\}$$

Note that $|\mathcal{M}'_1| = 2^{n(1-\mathbb{H}_b(\delta_2)-\epsilon_n)}$ and $p_{\bar{X}^n}(\bar{X}^n) = \frac{1}{|\mathcal{M}_1||\mathcal{M}'_1|}$. Therefore,

$$\begin{aligned} \mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{A}_0] &= \mathbb{P}_{\bar{X}^n \bar{Z}^n} \left[\frac{2^{-b\sqrt{n}}}{|\mathcal{M}_1|} < p_{\bar{X}^n|\bar{Z}^n}(\bar{X}^n|\bar{Z}^n) \leq \frac{2^{b\sqrt{n}}}{|\mathcal{M}_1|} \right] \\ &= \mathbb{P}_{\bar{X}^n \bar{Z}^n} \left[\frac{2^{-b\sqrt{n}}}{|\mathcal{M}_1|} < p_{\bar{Z}^n|\bar{X}^n}(\bar{Z}^n|\bar{X}^n) \frac{p_{\bar{X}^n}(\bar{X}^n)}{p_{\bar{Z}^n}(\bar{Z}^n)} \leq \frac{2^{b\sqrt{n}}}{|\mathcal{M}_1|} \right] \\ &= \mathbb{P}_{\bar{X}^n \bar{Z}^n} \left[\log \frac{p_{\bar{Z}^n|\bar{X}^n}(\bar{Z}^n|\bar{X}^n)}{p_{\bar{Z}^n}(\bar{Z}^n)} \leq b\sqrt{n} + n(1 - \mathbb{H}_b(\delta_2) - \epsilon_n) \right] \\ &\quad - \mathbb{P}_{\bar{X}^n \bar{Z}^n} \left[\log \frac{p_{\bar{Z}^n|\bar{X}^n}(\bar{Z}^n|\bar{X}^n)}{p_{\bar{Z}^n}(\bar{Z}^n)} \leq -b\sqrt{n} + n(1 - \mathbb{H}_b(\delta_2) - \epsilon_n) \right] \\ &= \mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{Q}_n^+] + \mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{Q}_n^-], \end{aligned}$$

where we have defined

$$\mathcal{Q}_n^\pm \triangleq \left\{ (x^n, z^n) \in \mathcal{C}_n \times \mathcal{Z}^n : \log \frac{p_{\bar{Z}^n|\bar{X}^n}(z^n|x^n)}{p_{\bar{Z}^n}(z^n)} \leq \pm b\sqrt{n} + n(1 - \mathbb{H}_b(\delta_2) - \epsilon_n) \right\}$$

We analyze $\mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{Q}_n^+]$ and $\mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{Q}_n^-]$ by introducing the sets

$$\begin{aligned} \mathcal{A}_n^\pm &\triangleq \left\{ (x^n, z^n) \in \mathcal{C}_n \times \mathcal{Z}^n : \log \frac{p_{\bar{Z}^n|\bar{X}^n}(z^n|x^n)}{p_{\bar{Z}^n}(z^n)} \leq \pm 2b\sqrt{n} + n(1 - \mathbb{H}_b(\delta_2) - \epsilon_n) \right\} \\ \mathcal{B}_n &\triangleq \left\{ (x^n, z^n) \in \mathcal{C}_n \times \mathcal{Z}^n : \log \frac{p_{\bar{Z}^n}(z^n)}{p_{\bar{Z}^n}(z^n)} < b\sqrt{n} \right\} \\ \text{and } \mathcal{D}_n &\triangleq \left\{ (x^n, z^n) \in \mathcal{C}_n \times \mathcal{Z}^n : \log \frac{p_{\bar{Z}^n}(z^n)}{p_{\bar{Z}^n}(z^n)} > -b\sqrt{n} \right\}. \end{aligned}$$

Using the law of total probability and the fact that $\mathcal{Q}_n^+ \cap \mathcal{B}_n \subset \mathcal{A}_n^+$, we now upper bound $\mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{Q}_n^+]$ as follows

$$\begin{aligned} \mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{Q}_n^+] &= \mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{Q}_n^+ \cap \mathcal{B}_n] + \mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{Q}_n^+ \cap \mathcal{B}_n^c] \\ &\leq \mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{A}_n^+] + \mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{B}_n^c] \end{aligned} \tag{22}$$

For n large enough,

$$\begin{aligned}
\mathbb{P}[\mathcal{B}_n^c] &= \frac{1}{b\sqrt{n}} \sum_{z^n \in \mathcal{Z}^n} b\sqrt{n} p_{\bar{Z}^n}(z^n) \mathbb{1} \left\{ \log \frac{p_{\bar{Z}^n}(z^n)}{p_{Z^n}(z^n)} \geq b\sqrt{n} \right\} \\
&\leq \frac{1}{b\sqrt{n}} \sum_{z^n \in \mathcal{Z}^n} p_{\bar{Z}^n}(z^n) \log \frac{p_{\bar{Z}^n}(z^n)}{p_{Z^n}(z^n)} \\
&= \frac{1}{b\sqrt{n}} \mathbb{D}(p_{\bar{Z}^n} \| p_{Z^n}) \\
&\stackrel{(a)}{\leq} \frac{1}{b\sqrt{n}} \mathbb{V}(p_{\bar{Z}^n}, p_{Z^n}) \log \frac{|\mathcal{Z}|^n}{\mathbb{V}(p_{\bar{Z}^n}, p_{Z^n})} \\
&\stackrel{(b)}{\leq} \frac{\sqrt{n}}{b} (\log |\mathcal{Z}| + \beta) 2^{-\beta n}
\end{aligned} \tag{23}$$

where (a) follows from [13, Lemma 1] and (b) follows from the fact that $x \mapsto x \log \frac{|\mathcal{Z}|^n}{x}$ is monotonously increasing for x small enough. In addition, note that

$$\mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{A}_n^+] = \mathbb{P}_{\bar{X}^n \bar{Z}^n} \left[\frac{1}{\sqrt{n}} \log \frac{p_{\bar{Z}^n | \bar{X}^n}(\bar{Z}^n | \bar{X}^n)}{p_{Z^n}(\bar{Z}^n)} \leq 2b + \sqrt{n}(1 - \mathbb{H}_b(\delta_2) - \epsilon_n) \right].$$

Since the noise is additive, we have $\bar{Z}_i = \bar{X}_i + E_i$ for $i \in \llbracket 1, n \rrbracket$ where $\{E_i\}_{i \geq 1}$ is i.i.d. with distribution $p_E \sim \mathcal{B}(\delta_2)$ and independent of $\{\bar{X}_i\}_{i \geq 1}$; hence, $p_{\bar{Z}^n | \bar{X}^n}(\bar{Z}^n | \bar{X}^n) = \prod_{i=1}^n p_E(E_i)$ and

$$\log \frac{p_{\bar{Z}^n | \bar{X}^n}(\bar{Z}^n | \bar{X}^n)}{p_{Z^n}(\bar{Z}^n)} = \sum_{i=1}^n \log \frac{p_E(E_i)}{1/2}.$$

The random variables in the sum are i.i.d. with mean $1 - \mathbb{H}_b(\delta_2)$, variance $\sigma > 0$, and third moment $\rho < \infty$. From the Berry-Esseen Theorem [44], there exists a universal constant $c > 0$ such that

$$\mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{A}_n^+] \leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{2b}{\sigma} - \frac{\sqrt{n}\epsilon_n}{\sigma}} e^{-\frac{x^2}{2}} dx + \frac{c}{\sqrt{n}} \frac{\rho}{\sigma^3}. \tag{24}$$

Similarly, using the law of total probability, the fact that $\mathcal{A}_n^- \cap \mathcal{D}_n \subset \mathcal{Q}_n^- \cap \mathcal{D}_n$ and the inclusion-exclusion principle, we lower bound $\mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{Q}_n^-]$ as follows

$$\begin{aligned}
\mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{Q}_n^-] &= \mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{Q}_n^- \cap \mathcal{D}_n] + \mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{Q}_n^- \cap \mathcal{D}_n^c] \\
&\geq \mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{A}_n^- \cap \mathcal{D}_n] \\
&\geq \mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{A}_n^-] + \mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{D}_n] - \mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{A}_n^- \cup \mathcal{D}_n] \\
&\geq \mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{A}_n^-] + \mathbb{P}_{\bar{X}^n \bar{Z}^n}[\mathcal{D}_n] - 1
\end{aligned} \tag{25}$$

Note that,

$$\begin{aligned}
\mathbb{P}[\mathcal{D}_n^c] &= \sum_{z^n \in \mathcal{Z}^n} p_{\bar{Z}^n}(z^n) \mathbb{1} \left\{ \log \frac{p_{\bar{Z}^n}(z^n)}{p_{Z^n}(z^n)} \leq -b\sqrt{n} \right\} \\
&\leq 2^{-b\sqrt{n}} \sum_{z^n \in \mathcal{Z}^n} p_{Z^n}(z^n) \\
&\leq 2^{-b\sqrt{n}}.
\end{aligned} \tag{26}$$

and, following the reasoning leading to (24),

$$\mathbb{P}_{\bar{X}^n \bar{Z}^n} [\mathcal{A}_n^-] \geq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\frac{2b}{\sigma} - \frac{\sqrt{n}\epsilon_n}{\sigma}} e^{-\frac{x^2}{2}} dx - \frac{c}{\sqrt{n}} \frac{\rho}{\sigma^3}. \tag{27}$$

Combining equations (22)-(27), we obtain

$$\begin{aligned}
\mathbb{P}_{\bar{X}^n \bar{Z}^n} [\mathcal{A}_0] &= \mathbb{P}_{\bar{X}^n \bar{Z}^n} [\mathcal{Q}_n^+] - \mathbb{P}_{\bar{X}^n \bar{Z}^n} [\mathcal{Q}_n^-] \\
&\leq \frac{1}{\sqrt{2\pi}} \int_{-\frac{2b}{\sigma} - \frac{\sqrt{n}\epsilon_n}{\sigma}}^{\frac{2b}{\sigma} - \frac{\sqrt{n}\epsilon_n}{\sigma}} e^{-\frac{x^2}{2}} dx + \frac{2c}{\sqrt{n}} \frac{\rho}{\sigma^3} + \frac{\sqrt{n}}{b} (\log |\mathcal{Z}| + \beta) 2^{-\beta n} + 2^{-b\sqrt{n}} \\
&\leq \frac{4b}{\sigma\sqrt{2\pi}} + \frac{2c}{\sqrt{n}} \frac{\rho}{\sigma^3} + \frac{\sqrt{n}}{b} (\log |\mathcal{Z}| + \beta) 2^{-\beta n} + 2^{-b\sqrt{n}}.
\end{aligned} \tag{28}$$

Combining (28) with (21), and using the assumption $\lim_{n \rightarrow \infty} \mathbb{P}_e^*(\mathcal{C}_n) = 0$ from (20), we have

$$\forall b > 0 \quad \lim_{n \rightarrow \infty} \mathbb{S}_2(\mathcal{C}_n) \geq 1 - \frac{4b}{\sigma\sqrt{2\pi}}.$$

Therefore, there exists $\eta > 0$ such that, for n large enough, $\mathbb{S}_2(\mathcal{C}_n) \geq \eta$. Notice that Proposition 1 immediately implies that there exists $\eta^* > 0$ such that $\lim_{n \rightarrow \infty} \mathbb{S}_1(\mathcal{C}_n) \geq \eta^*$.

APPENDIX D

LEMMAS USED IN THE ACHIEVABILITY PROOF OF THEOREM 2

The following notation is used throughout this appendix. We recall that U^n, X^n, Y^n, Z^n are the random variables defined by the random code generation with distribution given in (7). For any $(k, l, m) \in \llbracket 1, M_0 \rrbracket \times \llbracket 1, M_1 \rrbracket \times \llbracket 1, M'_1 \rrbracket$, the random variables representing the codewords u_k^n and x_{klm}^n obtained with the random code generation are denoted by U_k^n and X_{klm}^n .

The random variables that correspond to the use of a specific code \mathcal{C}_n are denoted by $\bar{U}^n, \bar{X}^n, \bar{Y}^n, \bar{Z}^n$ with distribution given by (4). The channel outputs that correspond to the transmission of u_k^n and x_{klm}^n are denoted by \bar{Y}_{klm}^n , and \bar{Z}_{klm}^n , respectively.

A. Proof of Lemma 1

By symmetry of the random code construction, we have

$$\begin{aligned}\mathbb{E}[\mathbb{P}_e(\mathcal{C}_n)] &= \frac{1}{M_0 M_1 M'_1} \sum_{k=1}^{M_0} \sum_{l=1}^{M_1} \sum_{m=1}^{M'_1} \mathbb{E}[\mathbb{P}_e(\mathcal{C}_n | M_0 = k, M_1 = l, M'_1 = m)] \\ &= \mathbb{E}[\mathbb{P}_e(\mathcal{C}_n | M_0 = 1, M_1 = 1, M'_1 = 1)],\end{aligned}$$

which we can analyze in terms of the events

$$\begin{aligned}E_1(k) &\triangleq \{(\bar{\mathbf{U}}_k^n, \bar{\mathbf{Y}}_{111}^n) \in \mathcal{T}_1^n | M_0 = 1, M_1 = 1, M'_1 = 1\} \\ E_2(k) &\triangleq \{(\bar{\mathbf{U}}_k^n, \bar{\mathbf{Z}}_{111}^n) \in \mathcal{T}_3^n | M_0 = 1, M_1 = 1, M'_1 = 1\} \\ E_3(k, l, m) &\triangleq \{(\bar{\mathbf{U}}_k^n, \bar{\mathbf{X}}_{klm}^n, \bar{\mathbf{Y}}_{111}^n) \in \mathcal{T}_2^n | M_0 = 1, M_1 = 1, M'_1 = 1\}.\end{aligned}$$

The average probability of error can then be written as

$$\mathbb{E}[\mathbb{P}_e(\mathcal{C}_n)] = \mathbb{E} \left[\mathbb{P} \left[E_1^c(1) \cup \bigcup_{k \neq 1} E_1(k) \cup E_2(k) \cup \bigcup_{k \neq 1} E_2(k) \cup E_3^c(1, 1, 1) \cup \bigcup_{(l, m) \neq (1, 1)} E_3(1, l, m) \right] \right],$$

where the expectation is with respect to $\{\mathbf{U}_k^n\}$ and $\{\mathbf{X}_{klm}^n\}$ for $(k, l, m) \in \llbracket 1, M_0 \rrbracket \times \llbracket 1, M_1 \rrbracket \times \llbracket 1, M'_1 \rrbracket$. It follows from standard arguments (see for instance [7, Chapter 3]) that $\mathbb{E}[\mathbb{P}_e(\mathcal{C}_n)] < \epsilon$ for n large enough provided

$$\begin{aligned}\frac{1}{n} \log M_0 &\leq \mathbf{p}\text{-}\liminf_{n \rightarrow \infty} \frac{1}{n} \mathbf{I}(\mathbf{U}^n; \mathbf{Y}^n) - 2\gamma \\ \frac{1}{n} \log M_0 &\leq \mathbf{p}\text{-}\liminf_{n \rightarrow \infty} \frac{1}{n} \mathbf{I}(\mathbf{U}^n; \mathbf{Z}^n) - 2\gamma \\ \frac{1}{n} \log M_1 M'_1 &\leq \mathbf{p}\text{-}\liminf_{n \rightarrow \infty} \frac{1}{n} \mathbf{I}(\mathbf{X}^n; \mathbf{Y}^n | \mathbf{U}^n) - 2\gamma.\end{aligned}\tag{29}$$

B. Proof of Lemma 2

We start by developing an upper bound for $\mathbb{S}_2(\mathcal{C}_n)$ that will be simpler to analyze. First, we have

$$\begin{aligned}\mathbb{S}_2(\mathcal{C}_n) &\triangleq \mathbb{V}(p_{\mathbf{M}_1 \bar{\mathbf{Z}}^n}, p_{\mathbf{M}_1} p_{\bar{\mathbf{Z}}^n}) \leq \mathbb{V}(p_{\bar{\mathbf{U}}^n \mathbf{M}_1 \bar{\mathbf{Z}}^n}, p_{\mathbf{M}_1} p_{\bar{\mathbf{U}}^n \bar{\mathbf{Z}}^n}) \\ &= \mathbb{E}_{\bar{\mathbf{U}}^n \mathbf{M}_1} \left[\mathbb{V}(p_{\bar{\mathbf{Z}}^n | \bar{\mathbf{U}}^n \mathbf{M}_1}, p_{\bar{\mathbf{Z}}^n | \bar{\mathbf{U}}^n}) \right].\end{aligned}$$

Next, we use Lemma 8 to further bound $\mathbb{S}_2(\mathcal{C}_n)$ as follows.

$$\begin{aligned}\mathbb{S}_2(\mathcal{C}_n) &\leq \mathbb{E}_{\bar{\mathbf{U}}^n \mathbf{M}_1} \left[\mathbb{V}(p_{\bar{\mathbf{Z}}^n | \bar{\mathbf{U}}^n \mathbf{M}_1}, p_{\mathbf{Z}^n | \mathbf{U}^n}) + \mathbb{V}(p_{\mathbf{Z}^n | \mathbf{U}^n}, p_{\bar{\mathbf{Z}}^n | \bar{\mathbf{U}}^n}) \right] \\ &\leq \mathbb{E}_{\bar{\mathbf{U}}^n \mathbf{M}_1} \left[\mathbb{V}(p_{\bar{\mathbf{Z}}^n | \bar{\mathbf{U}}^n \mathbf{M}_1}, p_{\mathbf{Z}^n | \mathbf{U}^n}) \right] + \mathbb{E}_{\bar{\mathbf{U}}^n} \left[\mathbb{V}(p_{\mathbf{Z}^n | \mathbf{U}^n}, p_{\bar{\mathbf{Z}}^n | \bar{\mathbf{U}}^n}) \right] \\ &\leq \mathbb{E}_{\bar{\mathbf{U}}^n \mathbf{M}_1} \left[\mathbb{V}(p_{\bar{\mathbf{Z}}^n | \bar{\mathbf{U}}^n \mathbf{M}_1}, p_{\mathbf{Z}^n | \mathbf{U}^n}) \right] + \mathbb{E}_{\bar{\mathbf{U}}^n} \left[\mathbb{V}(p_{\mathbf{M}_1} p_{\mathbf{Z}^n | \mathbf{U}^n}, p_{\bar{\mathbf{Z}}^n \mathbf{M}_1 | \bar{\mathbf{U}}^n}) \right] \\ &= 2\mathbb{E}_{\bar{\mathbf{U}}^n \mathbf{M}_1} \left[\mathbb{V}(p_{\bar{\mathbf{Z}}^n | \bar{\mathbf{U}}^n \mathbf{M}_1}, p_{\mathbf{Z}^n | \mathbf{U}^n}) \right].\end{aligned}\tag{30}$$

Notice that the term in brackets on the right hand side is a variational distance between the following two distributions:

- $p_{\bar{Z}^n|\bar{U}^n=\mathbf{u}_k^n, M_1=l}(z^n) = \sum_{m=1}^{M'_1} \frac{1}{M'_1} W_{Z^n|X^n}(z^n|x_{klm}^n)$, which represents the distribution induced at the eavesdropper's channel output by the M'_1 codewords $\{x_{kli}^n\}_{i \in \llbracket 1, M'_1 \rrbracket}$ selected with a uniform distribution.
- $p_{Z^n|U^n=\mathbf{u}_k^n}(z^n) = \sum_{x^n} W_{Z^n|X^n}(z^n|x^n) p_{X^n|U^n=\mathbf{u}_k^n}(x^n)$, which represents the distribution induced at the eavesdropper's channel output by an input process with distribution $p_{X^n|U^n=\mathbf{u}_k^n}(x^n)$.

Therefore, a *sufficient condition* for $\mathbb{S}_2(\mathcal{C}_n)$ to vanish is that, for every pair $(k, l) \in \llbracket 1, M_0 \rrbracket \times \llbracket 1, M_1 \rrbracket$, the variational distance between the two distributions vanishes as well. This is possible if each set of codewords $\{x_{kli}^n\}_{i \in \llbracket 1, M'_1 \rrbracket}$ approximates the same process with distribution $p_{Z^n|U^n=\mathbf{u}_k^n}(z^n)$ at the eavesdropper's output, which is exactly what the concept of channel resolvability reviewed in Section II is about. In other words, *a sufficient condition to guarantee secrecy is for each sub-codebook $\{x_{kli}^n\}_{i \in \llbracket 1, M'_1 \rrbracket}$ to be "resolvability code"*.

We establish the existence of such codebooks with a random coding argument following that used in [5]. The presence of a common message makes the proof slightly more involved but the steps remain essentially the same. On taking the average over \mathcal{C}_n for both sides of (30), we obtain

$$\mathbb{E}_{\mathcal{C}_n}[\mathbb{S}_2(\mathcal{C}_n)] \leq 2\mathbb{E}_{\bar{U}^n, M_1} \left[\mathbb{E}_{\mathcal{C}_n} \left[\mathbb{V} \left(p_{\bar{Z}^n|\bar{U}^n, M_1}, p_{Z^n|U^n} \right) \right] \right] \quad (31)$$

By symmetry of the random code construction, the inner expectation in (31) is the same for all values of $\bar{U}^n = \mathbf{u}_k^n$ and $M_1 = l$; hence, we have

$$\mathbb{E}_{\mathcal{C}_n}[\mathbb{S}_2(\mathcal{C}_n)] \leq 2\mathbb{E}_{\mathcal{C}_n} \left[\mathbb{V} \left(p_{\bar{Z}^n|\bar{U}^n=\mathbf{u}_1^n, M_1=1}, p_{Z^n|U^n=\mathbf{u}_1^n} \right) \right]. \quad (32)$$

Let $\tau > 0$. On using [7, Lemma 6.3.1] we finally upper bound (32) as

$$\begin{aligned} \mathbb{E}_{\mathcal{C}_n}[\mathbb{S}_2(\mathcal{C}_n)] &\leq 4\tau + 4A_n \\ \text{with } A_n &\triangleq \mathbb{E}_{\mathcal{C}_n} \left[\mathbb{P}_{\bar{Z}^n|\bar{U}^n=\mathbf{u}_1^n, M_1=1} \left[\log \frac{p_{\bar{Z}^n|\bar{U}^n=\mathbf{u}_1^n, M_1=1}(\bar{Z}^n)}{p_{Z^n|U^n=\mathbf{u}_1^n}(\bar{Z}^n)} > \tau \right] \right]. \end{aligned} \quad (33)$$

Note that the expectation over C_n reduces to the expectation over U_1^n and $\{X_{11j}\}_{j \in [1, M'_1]}$. Writing A_n explicitly, we obtain

$$\begin{aligned}
A_n &= \sum_{u_1^n \in \mathcal{U}^n} p_{U^n}(u_1^n) \sum_{x_{111} \in \mathcal{X}^n} p_{X^n|U^n}(x_{111}^n | u_1^n) \cdots \sum_{x_{11M'_1} \in \mathcal{X}^n} p_{X^n|U^n}(x_{11M'_1}^n | u_1^n) \\
&\quad \sum_{z^n \in \mathcal{Z}^n} p_{\bar{Z}^n|\bar{U}^n=u_1^n, M_1=1}(z^n) \mathbb{1} \left\{ \log \frac{p_{\bar{Z}^n|\bar{U}^n=u_1^n, M_1=1}(z^n)}{p_{Z^n|U^n}(z^n | u_1^n)} > \tau \right\} \\
&\stackrel{(a)}{=} \frac{1}{M'_1} \sum_{m=1}^{M'_1} \sum_{u_1^n \in \mathcal{U}^n} p_{U^n}(u_1^n) \sum_{x_{111} \in \mathcal{X}^n} p_{X^n|U^n}(x_{111}^n | u_1^n) \cdots \sum_{x_{11M'_1} \in \mathcal{X}^n} p_{X^n|U^n}(x_{11M'_1}^n | u_1^n) \\
&\quad \sum_{z^n \in \mathcal{Z}^n} W_{Z^n|X^n}(z^n | x_{11m}^n) \mathbb{1} \left\{ \log \frac{p_{\bar{Z}^n|\bar{U}^n=u_1^n, M_1=1}(z^n)}{p_{Z^n|U^n}(z^n | u_1^n)} > \tau \right\} \\
&\stackrel{(b)}{=} \sum_{u_1^n \in \mathcal{U}^n} p_{U^n}(u_1^n) \sum_{x_{112} \in \mathcal{X}^n} p_{X^n|U^n}(x_{112}^n | u_1^n) \cdots \sum_{x_{11M'_1} \in \mathcal{X}^n} p_{X^n|U^n}(x_{11M'_1}^n | u_1^n) \\
&\quad \sum_{x_{111} \in \mathcal{X}^n} \sum_{z^n \in \mathcal{Z}^n} W_{Z^n|X^n}(z^n | x_{111}^n) p_{X^n|U^n}(x_{111}^n | u_1^n) \mathbb{1} \left\{ \log \frac{p_{\bar{Z}^n|\bar{U}^n=u_1^n, M_1=1}(z^n)}{p_{Z^n|U^n}(z^n | u_1^n)} > \tau \right\}, \\
&\stackrel{(c)}{=} \sum_{u_1^n \in \mathcal{U}^n} p_{U^n}(u_1^n) \sum_{x_{112} \in \mathcal{X}^n} p_{X^n|U^n}(x_{112}^n | u_1^n) \cdots \sum_{x_{11M'_1} \in \mathcal{X}^n} p_{X^n|U^n}(x_{11M'_1}^n | u_1^n) \\
&\quad \sum_{x_{111} \in \mathcal{X}^n} \sum_{z^n \in \mathcal{Z}^n} p_{Z^n|X^n|U^n}(z^n, x_{111}^n | u_1^n) \mathbb{1} \left\{ \log \left(\frac{1}{M'_1} \sum_{m=1}^{M'_1} \frac{p_{Z^n|X^n|U^n}(z^n | x_{11m}^n u_1^n)}{p_{Z^n|U^n}(z^n | u_1^n)} \right) > \tau \right\} \quad (34)
\end{aligned}$$

where equality (a) follows from the definition of $p_{\bar{Z}^n|\bar{U}^n=u_1^n, M_1=1, C_n}(z^n)$, equality (b) follows by remarking that all codewords are generated according to the same density $p_{X^n|U^n}$ and equality (c) follows by noting that

- $W_{Z^n|X^n}(z^n | x_{111}^n) p_{X^n|U^n}(x_{111}^n | u_1^n) = p_{Z^n|X^n|U^n}(z^n, x_{111}^n | u_1^n)$ according to (7);
- for any u_1^n such that $p_{X^n|U^n}(x_{11m}^n | u_1^n) > 0$,

$$p_{\bar{Z}^n|\bar{U}^n=u_1^n, M_1=1}(z^n) = \frac{1}{M'_1} \sum_{m=1}^{M'_1} W_{Z^n|X^n}(z^n | x_{11m}^n) = \frac{1}{M'_1} \sum_{m=1}^{M'_1} p_{Z^n|X^n|U^n}(z^n | x_{11m}^n u_1^n).$$

Setting $\rho \triangleq \frac{2^\tau - 1}{2}$ we obtain

$$\begin{aligned}
&\log \left(\frac{1}{M'_1} \sum_{m=1}^{M'_1} \frac{p_{Z^n|X^n|U^n}(z^n | x_{11m}^n u_1^n)}{p_{Z^n|U^n}(z^n | u_1^n)} \right) > \tau \Leftrightarrow \\
&\frac{1}{M'_1} \exp \left(\log \frac{p_{Z^n|X^n|U^n}(z^n | x_{111}^n u_1^n)}{p_{Z^n|U^n}(z^n | u_1^n)} \right) + \frac{1}{M'_1} \sum_{m=2}^{M'_1} \exp \left(\log \frac{p_{Z^n|X^n|U^n}(z^n | x_{11m}^n u_1^n)}{p_{Z^n|U^n}(z^n | u_1^n)} \right) > 1 + 2\rho.
\end{aligned}$$

Therefore,

$$\begin{aligned}
A_n &\leq \mathbb{P}_{\mathbf{U}^n \mathbf{X}^n \mathbf{Z}^n} \left[\frac{1}{M'_1} \exp \left(\log \frac{p_{\mathbf{Z}^n | \mathbf{X}^n \mathbf{U}^n}(\mathbf{Z}^n | \mathbf{X}^n \mathbf{U}^n)}{p_{\mathbf{Z}^n | \mathbf{U}^n}(\mathbf{Z}^n | \mathbf{U}^n)} \right) > \rho \right] \\
&\quad + \mathbb{P}_{\mathbf{U}^n \mathbf{X}_{112}^n \dots \mathbf{X}_{11M'_1}^n \mathbf{Z}^n} \left[\frac{1}{M'_1} \sum_{j=2}^{M'_1} \exp \left(\log \frac{p_{\mathbf{Z}^n | \mathbf{X}^n \mathbf{U}^n}(\mathbf{z}^n | \mathbf{x}_{11j}^n \mathbf{u}_1^n)}{p_{\mathbf{Z}^n | \mathbf{U}^n}(\mathbf{z}^n | \mathbf{u}_1^n)} \right) > 1 + \rho \right] \\
&= \mathbb{P}_{\mathbf{U}^n \mathbf{X}^n \mathbf{Z}^n} \left[\frac{1}{n} \mathbf{I}(\mathbf{X}^n; \mathbf{Z}^n | \mathbf{U}^n) > \frac{1}{n} \log M'_1 + \frac{1}{n} \log \rho \right] \\
&\quad + \mathbb{E}_{\mathbf{U}^n \mathbf{Z}^n} \left[\mathbb{P}_{\mathbf{X}_{112}^n \dots \mathbf{X}_{11M'_1}^n | \mathbf{U}^n \mathbf{Z}^n} \left[\frac{1}{M'_1} \sum_{j=2}^{M'_1} \exp \left(\log \frac{p_{\mathbf{Z}^n | \mathbf{X}^n \mathbf{U}^n}(\mathbf{Z}^n | \mathbf{X}_{11j}^n \mathbf{U}^n)}{p_{\mathbf{Z}^n | \mathbf{U}^n}(\mathbf{Z}^n | \mathbf{U}^n)} \right) > 1 + \rho \right] \right]. \tag{35}
\end{aligned}$$

By (34), note that, conditioned on $\mathbf{U}^n = \mathbf{u}^n$, $\{\mathbf{X}_{11j}\}_{j \in \llbracket 2, M'_1 \rrbracket}$ are i.i.d with distribution $p_{\mathbf{X}^n | \mathbf{U}^n = \mathbf{u}^n}$ and independent of \mathbf{Z}^n . To analyze the second term on the right-hand side, let us define

$$B_n(\mathbf{z}^n, \mathbf{u}^n) \triangleq \mathbb{P}_{\mathbf{X}_{112}^n \dots \mathbf{X}_{11M'_1}^n | \mathbf{U}^n = \mathbf{u}^n} \left[\frac{1}{M'_1} \sum_{j=2}^{M'_1} \exp \left(\log \frac{p_{\mathbf{Z}^n | \mathbf{X}^n \mathbf{U}^n}(\mathbf{z}^n | \mathbf{X}_{11j}^n \mathbf{u}^n)}{p_{\mathbf{Z}^n | \mathbf{U}^n}(\mathbf{z}^n | \mathbf{u}^n)} \right) > 1 + \rho \right].$$

Let us introduce the random variables

$$\begin{aligned}
D_j^n(\mathbf{z}^n, \mathbf{u}^n) &\triangleq \exp \left(\log \frac{p_{\mathbf{Z}^n | \mathbf{X}^n \mathbf{U}^n}(\mathbf{z}^n | \mathbf{X}_{11j}^n \mathbf{u}^n)}{p_{\mathbf{Z}^n | \mathbf{U}^n}(\mathbf{z}^n | \mathbf{u}^n)} \right) \quad \text{for } j \in \llbracket 2, M'_1 \rrbracket \\
E_j^n(\mathbf{z}^n, \mathbf{u}^n) &\triangleq D_j^n(\mathbf{z}^n, \mathbf{u}^n) \mathbf{1} \{D_j^n(\mathbf{z}^n, \mathbf{u}^n) \leq M'_1\} \quad \text{for } j \in \llbracket 2, M'_1 \rrbracket \\
F_{M'_1}(\mathbf{z}^n, \mathbf{u}^n) &\triangleq \frac{1}{M'_1} \sum_{j=2}^{M'_1} D_j^n(\mathbf{z}^n, \mathbf{u}^n) \\
G_{M'_1}(\mathbf{z}^n, \mathbf{u}^n) &\triangleq \frac{1}{M'_1} \sum_{j=2}^{M'_1} E_j^n(\mathbf{z}^n, \mathbf{u}^n).
\end{aligned}$$

For a fixed $\mathbf{z}^n, \mathbf{u}^n$, the random variables $\{D_j^n(\mathbf{z}^n, \mathbf{u}^n)\}_{j \in \llbracket 2, M'_1 \rrbracket}$ are i.i.d. by construction, and so are the random variables $\{E_j^n(\mathbf{z}^n, \mathbf{u}^n)\}_{j \in \llbracket 2, M'_1 \rrbracket}$. By the law of total probability,

$$\begin{aligned}
B_n(\mathbf{z}^n, \mathbf{u}^n) &= \mathbb{P}_{\mathbf{X}_{112}^n \dots \mathbf{X}_{11M'_1}^n | \mathbf{U}^n = \mathbf{u}^n} [F_{M'_1}(\mathbf{z}^n, \mathbf{u}^n) > 1 + \rho] \\
&\leq \underbrace{\mathbb{P}_{\mathbf{X}_{112}^n \dots \mathbf{X}_{11M'_1}^n | \mathbf{U}^n = \mathbf{u}^n} [G_{M'_1}(\mathbf{z}^n, \mathbf{u}^n) > 1 + \rho]}_{\triangleq C_n(\mathbf{z}^n, \mathbf{u}^n)} \\
&\quad + \underbrace{\mathbb{P}_{\mathbf{X}_{112}^n \dots \mathbf{X}_{11M'_1}^n | \mathbf{U}^n = \mathbf{u}^n} [G_{M'_1}(\mathbf{z}^n, \mathbf{u}^n) \neq F_{M'_1}(\mathbf{z}^n, \mathbf{u}^n)]}_{\triangleq D_n(\mathbf{z}^n, \mathbf{u}^n)}. \tag{36}
\end{aligned}$$

We first bound $D_n(z^n, u^n)$ as follows. Note that

$$\begin{aligned}
D_n(z^n, u^n) &\leq \mathbb{P}_{X_{112}^n \dots X_{11M'_1}^n | U^n = u^n} \left[\bigcup_{j=2}^{M'_1} \{E_j^n(z^n, u^n) \neq D_j^n(z^n, u^n)\} \right] \\
&\leq \sum_{j=2}^{M'_1} \mathbb{P}_{X_{11j}^n | U^n = u^n} [E_j^n(z^n, u^n) \neq D_j^n(z^n, u^n)] \\
&\leq M'_1 \mathbb{P}_{X_{112}^n | U^n = u^n} [D_2^n(z^n, u^n) > M'_1].
\end{aligned}$$

Therefore, on taking the expectation over $Z^n U^n$, we have

$$\begin{aligned}
&\mathbb{E}_{Z^n U^n} [D_n(Z^n, U^n)] \\
&\leq M'_1 \sum_{z^n} \sum_{u^n} p_{Z^n U^n}(z^n, u^n) \mathbb{P}_{X_{112}^n | U^n = u^n} [D_2^n(z^n, u^n) > M'_1] \\
&= M'_1 \sum_{z^n} \sum_{u^n} \sum_{x^n} p_{Z^n U^n}(z^n, u^n) p_{X^n | U^n}(x^n | u^n) \mathbb{1} \left\{ \exp \left(\log \frac{p_{Z^n | X^n U^n}(z^n | x^n, u^n)}{p_{Z^n | U^n}(z^n | u^n)} \right) > M'_1 \right\} \\
&\leq \sum_{z^n} \sum_{u^n} \sum_{x^n} p_{Z^n U^n}(z^n, u^n) p_{X^n | U^n}(x^n | u^n) \\
&\quad \exp \left(\log \frac{p_{Z^n | X^n U^n}(z^n | x^n, u^n)}{p_{Z^n | U^n}(z^n | u^n)} \right) \mathbb{1} \left\{ \exp \left(\log \frac{p_{Z^n | X^n U^n}(z^n | x^n, u^n)}{p_{Z^n | U^n}(z^n | u^n)} \right) > M'_1 \right\} \\
&= \sum_{z^n} \sum_{u^n} \sum_{x^n} p_{Z^n X^n U^n}(z^n, x^n, u^n) \mathbb{1} \left\{ \exp \left(\log \frac{p_{Z^n | X^n U^n}(z^n | x^n, u^n)}{p_{Z^n | U^n}(z^n | u^n)} \right) > M'_1 \right\} \\
&= \mathbb{P}_{Z^n X^n U^n} \left[\frac{1}{n} \log \frac{p_{Z^n | X^n U^n}(Z^n | X^n, U^n)}{p_{Z^n | U^n}(Z^n | U^n)} > \frac{1}{n} \log M'_1 \right]. \tag{37}
\end{aligned}$$

Finally, we bound $\mathbb{E}_{Z^n U^n} [C_n(Z^n, U^n)]$ using Chebyshev's inequality. Note that

$$\begin{aligned}
&\mathbb{E}_{X_{112}^n \dots X_{11M'_1}^n | U^n = u^n} [G_{M'_1}(z^n, u^n)] \\
&= \frac{1}{M'_1} \sum_{j=2}^{M'_1} \mathbb{E}_{X_{11j}^n | U^n = u^n} [E_j^n(z^n, u^n)] \\
&= \mathbb{E}_{X_{112}^n | U^n = u^n} [E_2^n(z^n, u^n)] \\
&= \sum_{x^n} p_{X^n | U^n}(x^n | u^n) \exp \left(\log \frac{p_{Z^n | X^n U^n}(z^n | x^n, u^n)}{p_{Z^n | U^n}(z^n | u^n)} \right) \mathbb{1} \left\{ \exp \left(\log \frac{p_{Z^n | X^n U^n}(z^n | x^n, u^n)}{p_{Z^n | U^n}(z^n | u^n)} \right) \leq M'_1 \right\} \\
&= \sum_{x^n} p_{X^n | Z^n U^n}(x^n | z^n, u^n) \mathbb{1} \left\{ \exp \left(\log \frac{p_{Z^n | X^n U^n}(z^n | x^n, u^n)}{p_{Z^n | U^n}(z^n | u^n)} \right) \leq M'_1 \right\} \\
&\leq 1,
\end{aligned}$$

where the last equality follows from Bayes' rule. Define

$$\text{Var}(G_{M'_1}(z^n, u^n)) \triangleq \mathbb{E}_{X_{112}^n | U^n = u^n} \left[(G_{M'_1}(z^n, u^n) - \mathbb{E}_{X_{112}^n | U^n = u^n} [G_{M'_1}(z^n, u^n)])^2 \right].$$

Note that

$$\text{Var}(G_{M'_1}(z^n, u^n)) = \frac{1}{M'_1} \text{Var}(E_2^n(z^n, u^n));$$

hence, on applying Chebyshev's inequality, we have

$$\begin{aligned} \mathbb{P}_{X_{112}^n \dots X_{11M'_1}^n | U^n = u^n} [G_{M'_1}(z^n, u^n) > 1 + \rho] \\ &\leq \mathbb{P}_{X_{112}^n \dots X_{11M'_1}^n | U^n = u^n} [G_{M'_1}(z^n, u^n) - \mathbb{E}_{X_{112}^n \dots X_{11M'_1}^n | U^n = u^n} [G_{M'_1}(z^n, u^n)] > \rho] \\ &\leq \frac{1}{\rho^2} \text{Var}(G_{M'_1}(z^n, u^n)) \\ &= \frac{1}{\rho^2 M'_1} \text{Var}(E_2^n(z^n, u^n)) \end{aligned}$$

Therefore, on taking the expectation over $Z^n U^n$, we obtain

$$\begin{aligned} \mathbb{E}_{Z^n U^n} [C_n(Z^n, U^n)] &= \sum_{u^n} \sum_{z^n} p_{Z^n U^n}(z^n, u^n) \mathbb{P}_{X_{112}^n \dots X_{11M'_1}^n | U^n = u^n} [G_{M'_1}(z^n, u^n) > 1 + \rho] \\ &\leq \sum_{u^n} \sum_{z^n} p_{Z^n U^n}(z^n, u^n) \frac{1}{\rho^2 M'_1} \text{Var}(E_2^n(z^n, u^n)) \\ &= \frac{1}{\rho^2 M'_1} \text{Var}(E_2^n(Z^n, U^n)) \\ &\leq \frac{1}{\rho^2 M'_1} \mathbb{E}_{Z^n U^n} [\mathbb{E}_{X_{112}^n | U^n} [E_2^n(Z^n, U^n)^2]] \end{aligned} \quad (38)$$

Finally, note that

$$\begin{aligned} &\frac{1}{M'_1} \mathbb{E}_{Z^n U^n} [\mathbb{E}_{X_{112}^n | U^n} [E_2^n(Z^n, U^n)^2]] \\ &= \frac{1}{M'_1} \sum_{u^n} \sum_{x^n} \sum_{z^n} p_{Z^n U^n}(z^n, u^n) p_{X^n | U^n}(x^n | u^n) \left(\frac{p_{Z^n | X^n U^n}(z^n | x^n, u^n)}{p_{Z^n | U^n}(z^n | u^n)} \right)^2 \mathbb{1} \left\{ \frac{p_{Z^n | X^n U^n}(z^n | x^n, u^n)}{p_{Z^n | U^n}(z^n | u^n)} \leq M'_1 \right\} \\ &= \frac{1}{M'_1} \sum_{u^n} \sum_{x^n} \sum_{z^n} p_{Z^n X^n U^n}(z^n, x^n, u^n) \frac{p_{Z^n | X^n U^n}(z^n | x^n, u^n)}{p_{Z^n | U^n}(z^n | u^n)} \mathbb{1} \left\{ \frac{p_{Z^n | X^n U^n}(z^n | x^n, u^n)}{p_{Z^n | U^n}(z^n | u^n)} \leq M'_1 \right\} \\ &= \mathbb{E}_{U^n X^n Z^n} \left[\frac{1}{M'_1} \exp I(X^n; Z^n | U^n) \mathbb{1} \left\{ \exp I(X^n; Z^n | U^n) \leq M'_1 \right\} \right] \\ &\leq 2^{-n\gamma} + \mathbb{P}_{U^n X^n Z^n} \left[\frac{1}{n} I(X^n; Z^n | U^n) \geq \frac{\log M'_1}{n} - \gamma \right]. \end{aligned} \quad (39)$$

Combining (35), (36) (37), (38) and (39), we obtain that for any $\tau > 0$

$$\begin{aligned} \mathbb{E}_{C_n}[\mathbb{S}_2(C_n)] &\leq 4\tau + 4\mathbb{P}_{U^n X^n Z^n} \left[\frac{1}{n} I(X^n; Z^n | U^n) \geq \frac{\log M'_1}{n} + \frac{\log \rho}{n} \right] \\ &\quad + 4\mathbb{P}_{U^n X^n Z^n} \left[\frac{1}{n} I(X^n; Z^n | U^n) \geq \frac{\log M'_1}{n} \right] \\ &\quad + \frac{4 \cdot 2^{-n\gamma}}{\rho^2} + \frac{4}{\rho^2} \mathbb{P}_{U^n X^n Z^n} \left[\frac{1}{n} I(X^n; Z^n | U^n) \geq \frac{\log M'_1}{n} - \gamma \right]. \end{aligned} \quad (40)$$

Therefore, $\mathbb{E}_{C_n}[\mathbb{S}_2(C_n)] < \epsilon$ for n large enough provided

$$\frac{1}{n} \log M'_1 \geq \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(X^n; Z^n | U^n) + 2\gamma. \quad (41)$$

APPENDIX E

LEMMAS USED IN THE CONVERSE PROOF OF THEOREM 2

A. Proof of Lemma 3

Following the proof of the Verdú-Han Lemma, one can easily show that

$$3\mathbb{P}_e(C_n) \geq \mathbb{P} \left[\begin{array}{l} \frac{1}{n} I(\bar{U}^n; \bar{Y}^n) \leq R_0 - \gamma \\ \text{or } \frac{1}{n} I(\bar{U}^n; \bar{Z}^n) \leq R_0 - \gamma \\ \text{or } \frac{1}{n} I(\bar{W}^n; \bar{Y}^n | \bar{U}^n) \leq R_1 - \gamma \end{array} \right] - 3 \cdot 2^{-n\gamma}, \quad (42)$$

from which the lemma follows.

B. Proof of Lemma 4

To prove Lemma 4, note that, with probability one,

$$\begin{aligned} \frac{1}{n} I(\bar{W}^n; \bar{Z}^n) &= \frac{1}{n} I(\bar{W}^n; \bar{Z}^n \bar{U}^n) - \frac{1}{n} I(\bar{W}^n; \bar{U}^n | \bar{Z}^n) \\ &= \frac{1}{n} I(\bar{W}^n; \bar{Z}^n | \bar{U}^n) - \frac{1}{n} I(\bar{W}^n; \bar{U}^n | \bar{Z}^n) \\ &= \frac{1}{n} I(\bar{W}^n; \bar{Z}^n | \bar{U}^n) - \frac{1}{n} H(\bar{U}^n | \bar{Z}^n) + \frac{1}{n} H(\bar{U}^n | \bar{W}^n \bar{Z}^n), \end{aligned}$$

where the second inequality follow from the independence of \bar{W}^n and \bar{U}^n . Consequently,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{S}_6(C_n) &= \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(\bar{W}^n; \bar{Z}^n) \\ &\geq \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(\bar{W}^n; \bar{Z}^n | \bar{U}^n) - \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} H(\bar{U}^n | \bar{Z}^n) + \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(\bar{U}^n | \bar{W}^n \bar{Z}^n). \end{aligned}$$

Note that $\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(\bar{U}^n | \bar{W}^n \bar{Z}^n) \geq 0$ and that (42) implies

$$3\mathbb{P}_e(\mathcal{C}_n) \geq \mathbb{P}\left[\frac{1}{n} H(\bar{U}^n | \bar{Z}^n) \geq \gamma\right] - 3 \cdot 2^{-n\gamma}.$$

Since $\lim_{n \rightarrow \infty} \mathbb{P}_e(\mathcal{C}_n) = 0$ and $\gamma > 0$ can be chosen arbitrarily small, we have $\text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} H(\bar{U}^n | \bar{Z}^n) = 0$. As $\lim_{n \rightarrow \infty} \mathbb{S}_6(\mathcal{C}_n) = 0$, we finally obtain

$$\text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(\bar{W}^n; \bar{Z}^n | \bar{U}^n) = 0.$$

APPENDIX F

PROOF OF THEOREM 3

We prove Theorem 3 with minor modifications of the proof of Theorem 2. Specifically, we establish secrecy for \mathbb{S}_1 by showing that there exist sequences of codes $\{\mathcal{C}_n\}_{n \geq 1}$ for which $\mathbb{S}_2(\mathcal{C}_n)$ decreases exponentially fast with n and by using [13, Lemma 1] to obtain an upper bound for $\mathbb{S}_1(\mathcal{C}_n)$. We handle the power constraint by using an appropriate distribution during the random code generation process as in [7, Section 3.2]. We note that a similar proof has been used by He and Yener in [45].

Let $\gamma, \delta, \epsilon > 0$. Let \mathcal{U} be an arbitrary discrete alphabet and fix a distribution $p_{\tilde{U}}$ on \mathcal{U} . Fix a conditional distribution $p_{\tilde{X}|\tilde{U}}$ on $\mathcal{X} \times \mathcal{U}$ such that $\mathbb{E}[c(\tilde{X})] \leq P - \delta$. Let $\tilde{U}^n, \tilde{X}^n, \tilde{Z}^n$ be the random variables with joint distribution

$$\forall (z^n, x^n, u^n) \in \mathcal{Z}^n \times \mathcal{X}^n \times \mathcal{U}^n \quad p_{\tilde{Z}^n \tilde{X}^n \tilde{U}^n}(z^n, x^n, u^n) = \prod_{i=1}^n W_{Z|X}(z_i | x_i) p_{\tilde{X}|\tilde{U}}(x_i | u_i) p_{\tilde{U}}(u_i).$$

We assume that $\tilde{U}^n, \tilde{X}^n, \tilde{Z}^n$ are such that the moment generating functions of $c(\tilde{X})$ and $I(\tilde{X}; \tilde{Z} | \tilde{U})$ converge unconditionally in a neighborhood of 0 and are differentiable at 0.

Define the set \mathcal{P}_n as

$$\mathcal{P}_n \triangleq \left\{ x^n \in \mathcal{X}^n : \frac{1}{n} \sum_{i=1}^n c(x_i) \leq P \right\}.$$

Lemma 7 shows that there exists $\alpha_\delta > 0$ such that $\mathbb{P}[\tilde{X}^n \in \mathcal{P}_n] \geq 1 - 2^{-\alpha_\delta n}$. In the sequel, we define $\gamma_n \triangleq 1 - 2^{-n \frac{\alpha_\delta}{2}}$. Define the set $\mathcal{G}_n \subset \mathcal{U}^n$ as follows:

$$\mathcal{G}_n \triangleq \left\{ u^n : \mathbb{P}_{\tilde{X}^n | \tilde{U}^n = u^n} [\tilde{X}^n \notin \mathcal{P}_n | \tilde{U}^n = u^n] < 2^{-n \frac{\alpha_\delta}{2}} \right\}.$$

Upon using Markov's inequality, we obtain

$$\begin{aligned}
\mathbb{P}_{\tilde{\mathbf{U}}^n} [\tilde{\mathbf{U}}^n \notin \mathcal{G}_n] &= \mathbb{P}_{\tilde{\mathbf{U}}^n} \left[\mathbb{P}_{\tilde{\mathbf{X}}^n | \tilde{\mathbf{U}}^n} [\tilde{\mathbf{X}}^n \notin \mathcal{P}_n | \tilde{\mathbf{U}}^n] \geq 2^{-n \frac{\alpha_\delta}{2}} \right] \\
&\leq \mathbb{E}_{\tilde{\mathbf{U}}^n} \left[\mathbb{P}_{\tilde{\mathbf{X}}^n | \tilde{\mathbf{U}}^n} [\tilde{\mathbf{X}}^n \notin \mathcal{P}_n | \tilde{\mathbf{U}}^n] \right] 2^{n \frac{\alpha_\delta}{2}} \\
&= \mathbb{P}_{\tilde{\mathbf{X}}^n} [\tilde{\mathbf{X}}^n \notin \mathcal{P}_n] 2^{n \frac{\alpha_\delta}{2}} \\
&\leq 2^{-n(\alpha_\delta - \frac{\alpha_\delta}{2})} \\
&= 2^{-n \frac{\alpha_\delta}{2}}.
\end{aligned}$$

Now, we define the random variables $\mathbf{U}^n, \mathbf{X}^n, \mathbf{Z}^n$ as follows. First,

$$\forall \mathbf{u}^n \in \mathcal{U}^n \quad p_{\mathbf{U}^n}(\mathbf{u}^n) = \begin{cases} \frac{1}{\mathbb{P}_{\tilde{\mathbf{U}}^n}[\tilde{\mathbf{U}}^n \in \mathcal{G}_n]} p_{\tilde{\mathbf{U}}^n}(\mathbf{u}^n) & \text{if } \mathbf{u}^n \in \mathcal{G}_n \\ 0 & \text{else.} \end{cases}$$

By construction, we have

$$\forall \mathbf{u}^n \in \mathcal{U}^n \quad p_{\mathbf{U}^n}(\mathbf{u}^n) \leq \frac{p_{\tilde{\mathbf{U}}^n}(\mathbf{u}^n)}{\gamma_n}. \quad (43)$$

Next,

$$\forall (\mathbf{x}^n, \mathbf{u}^n) \in \mathcal{X}^n \times \mathcal{G}_n \quad p_{\mathbf{X}^n | \mathbf{U}^n}(\mathbf{x}^n | \mathbf{u}^n) = \begin{cases} \frac{1}{\mathbb{P}_{\tilde{\mathbf{X}}^n | \tilde{\mathbf{U}}^n = \mathbf{u}^n}[\tilde{\mathbf{X}}^n \in \mathcal{P}_n | \tilde{\mathbf{U}}^n = \mathbf{u}^n]} p_{\tilde{\mathbf{X}}^n | \tilde{\mathbf{U}}^n}(\mathbf{x}^n | \mathbf{u}^n) & \text{if } \mathbf{x}^n \in \mathcal{P}_n \\ 0 & \text{else.} \end{cases}$$

By construction, we have

$$\forall (\mathbf{x}^n, \mathbf{u}^n) \in \mathcal{X}^n \times \mathcal{G}_n \quad p_{\mathbf{X}^n | \mathbf{U}^n}(\mathbf{x}^n | \mathbf{u}^n) \leq \frac{p_{\tilde{\mathbf{X}}^n | \tilde{\mathbf{U}}^n}(\mathbf{x}^n | \mathbf{u}^n)}{\gamma_n}. \quad (44)$$

Finally,

$$\forall (\mathbf{z}^n, \mathbf{x}^n, \mathbf{u}^n) \in \mathcal{Z}^n \times \mathcal{X}^n \times \mathcal{G}_n \quad p_{\mathbf{Z}^n | \mathbf{X}^n \mathbf{U}^n}(\mathbf{z}^n, \mathbf{x}^n, \mathbf{u}^n) = W_{\mathbf{Z}^n | \mathbf{X}^n}(\mathbf{z}^n | \mathbf{x}^n) p_{\mathbf{X}^n | \mathbf{U}^n}(\mathbf{x}^n | \mathbf{u}^n) p_{\mathbf{U}^n}(\mathbf{u}^n). \quad (45)$$

We repeat the random coding argument in the proof of Theorem 2 with the distribution $p_{\mathbf{X}^n \mathbf{U}^n}$ defined by (45).

Lemma 12 (Reliability conditions).

$$\begin{cases} R_0 \leq \min \left(\mathbb{I}(\tilde{\mathbf{U}}; \tilde{\mathbf{Y}}) - 2\gamma, \mathbb{I}(\tilde{\mathbf{U}}; \tilde{\mathbf{Z}}) - 2\gamma \right) \\ R_1 + R'_1 \leq \mathbb{I}(\tilde{\mathbf{X}}; \tilde{\mathbf{Y}} | \tilde{\mathbf{U}}) - 2\gamma, \end{cases} \Rightarrow \lim_{n \rightarrow \infty} \mathbb{E}[\mathbb{P}_e(\mathbf{C}_n)] \leq \epsilon.$$

Proof: Following [7, Proof of Theorem 3.6.2], one can show that

$$\begin{aligned} \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{U}^n; \mathbf{Y}^n) &\geq \mathbb{I}(\tilde{\mathbf{U}}; \tilde{\mathbf{Y}}), & \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{U}^n; \mathbf{Z}^n) &\geq \mathbb{I}(\tilde{\mathbf{U}}; \tilde{\mathbf{Z}}) \\ \text{and } \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{X}^n; \mathbf{Y}^n | \mathbf{U}^n) &\geq \mathbb{I}(\tilde{\mathbf{X}}; \tilde{\mathbf{Y}} | \tilde{\mathbf{U}}). \end{aligned}$$

Hence, the result follows directly from Lemma 1. ■

Lemma 13 (Secrecy from resolvability conditions). *There exists $\alpha_{\delta, \gamma} > 0$, such that*

$$R'_1 \geq \mathbb{I}(\tilde{\mathbf{X}}; \tilde{\mathbf{Z}} | \tilde{\mathbf{U}}) + 2\gamma, \Rightarrow \lim_{n \rightarrow \infty} \mathbb{E}[\mathbb{P}_e(C_n)] \leq 2^{-\alpha_{\delta, \gamma}}.$$

Proof: Note that (30) still holds. Upon using Lemma 8, we obtain

$$\begin{aligned} \mathbb{E}_{C_n}[\mathbb{S}_2(C_n)] &\leq 2\mathbb{E}_{C_n} \left[\mathbb{V} \left(p_{\tilde{\mathbf{Z}}^n | \tilde{\mathbf{U}}^n = \mathbf{U}_1^n, M_1=1, C_n}, p_{\mathbf{Z}^n | \mathbf{U}^n = \mathbf{U}_1^n} \right) \right] \\ &\leq 2\mathbb{E}_{C_n} \left[\mathbb{V} \left(p_{\tilde{\mathbf{Z}}^n | \tilde{\mathbf{U}}^n = \mathbf{U}_1^n, M_1=1, C_n}, p_{\tilde{\mathbf{Z}}^n | \tilde{\mathbf{U}}^n = \mathbf{U}_1^n} \right) \right] + 2\mathbb{E}_{C_n} \left[\mathbb{V} \left(p_{\tilde{\mathbf{Z}}^n | \tilde{\mathbf{U}}^n = \mathbf{U}_1^n}, p_{\mathbf{Z}^n | \mathbf{U}^n = \mathbf{U}_1^n} \right) \right]. \end{aligned} \quad (46)$$

First, we bound the second term on the right-hand side of (46). For all $\mathbf{u}_1^n \in \mathcal{G}_n$,

$$\begin{aligned} &\mathbb{V} \left(p_{\tilde{\mathbf{Z}}^n | \tilde{\mathbf{U}}^n = \mathbf{u}_1^n}, p_{\mathbf{Z}^n | \mathbf{U}^n = \mathbf{u}_1^n} \right) \\ &\stackrel{(a)}{\leq} \mathbb{V} \left(p_{\tilde{\mathbf{X}}^n | \tilde{\mathbf{U}}^n = \mathbf{u}_1^n}, p_{\mathbf{X}^n | \mathbf{U}^n = \mathbf{u}_1^n} \right) \\ &= 2 \sup_{\mathcal{A} \subseteq \mathcal{X}^n} \left| \mathbb{P}_{\mathbf{X}^n | \mathbf{U}^n = \mathbf{u}_1^n}[\mathcal{A}] - \mathbb{P}_{\tilde{\mathbf{X}}^n | \tilde{\mathbf{U}}^n = \mathbf{u}_1^n}[\mathcal{A}] \right| \\ &= \sup_{\mathcal{A} \subseteq \mathcal{X}^n} \sum_{\mathcal{B} \in \{\mathcal{A}, \mathcal{A}^c\}} \left(\left| \mathbb{P}_{\mathbf{X}^n | \mathbf{U}^n = \mathbf{u}_1^n}[\mathcal{B}] - \mathbb{P}_{\tilde{\mathbf{X}}^n | \tilde{\mathbf{U}}^n = \mathbf{u}_1^n}[\mathcal{B}] \right| \right) \\ &= \sup_{\mathcal{A} \subseteq \mathcal{X}^n} \sum_{\mathcal{B} \in \{\mathcal{A}, \mathcal{A}^c\}} \left(\left| \mathbb{P}_{\mathbf{X}^n | \mathbf{U}^n = \mathbf{u}_1^n}[\mathcal{B} \cap \mathcal{P}_n] - \mathbb{P}_{\tilde{\mathbf{X}}^n | \tilde{\mathbf{U}}^n = \mathbf{u}_1^n}[\mathcal{B} \cap \mathcal{P}_n^c] - \mathbb{P}_{\tilde{\mathbf{X}}^n | \tilde{\mathbf{U}}^n = \mathbf{u}_1^n}[\mathcal{B} \cap \mathcal{P}_n] \right| \right) \\ &\stackrel{(b)}{\leq} \sup_{\mathcal{A} \subseteq \mathcal{X}^n} \sum_{\mathcal{B} \in \{\mathcal{A}, \mathcal{A}^c\}} \left(\mathbb{P}_{\tilde{\mathbf{X}}^n | \tilde{\mathbf{U}}^n = \mathbf{u}_1^n}[\mathcal{B} \cap \mathcal{P}_n] \left(\frac{1}{\gamma_n} - 1 \right) + \mathbb{P}_{\tilde{\mathbf{X}}^n | \tilde{\mathbf{U}}^n = \mathbf{u}_1^n}[\mathcal{B} \cap \mathcal{P}_n^c] \right) \\ &= \left(\frac{1}{\gamma_n} - 1 \right) + (1 - \gamma_n), \end{aligned}$$

where (a) follows from Lemma 9 and (b) follows from the definition of $p_{\mathbf{X}^n | \mathbf{U}^n}$ in (45) and the bound in (44); therefore, for n large enough there exists $\beta_\delta > 0$, such that

$$\mathbb{E}_{C_n} \left[\mathbb{V} \left(p_{\tilde{\mathbf{Z}}^n | \tilde{\mathbf{U}}^n = \mathbf{U}_1^n}, p_{\mathbf{Z}^n | \mathbf{U}^n = \mathbf{U}_1^n} \right) \right] \leq 2^{-\beta_\delta n} \quad (47)$$

We now bound the first term on the right-hand side of (46). Applying [7, Lemma 6.3.1], we obtain

$$2\mathbb{E}_{C_n} \left[\mathbb{V} \left(p_{\tilde{Z}^n | \tilde{U}^n = U_1^n, M_1=1}, p_{\tilde{Z}_1^n | \tilde{U}^n = U_1^n} \right) \right] \leq 4\tau + 4\mathbb{P}_{\tilde{Z}^n | \tilde{U}^n = U_1^n, M_1=1} \left[\log \frac{p_{\tilde{Z}^n | \tilde{U}^n = U_1^n, M_1=1}(\tilde{Z}^n)}{p_{\tilde{Z}^n | \tilde{U}^n = U_1^n}(\tilde{Z}^n)} > \tau \right]. \quad (48)$$

Note that (48) is similar to (33), and the only difference is the presence of $p_{\tilde{Z} | \tilde{U}}$ instead of $p_{Z | U}$ in the denominator; using the definition of $p_{Z^n X^n U^n}$ in (45), the bounds in (43) and (44), and repeating the steps leading from (33) to (40), one obtains after some calculations

$$\begin{aligned} 2\mathbb{E}_{C_n} \left[\mathbb{V} \left(p_{\tilde{Z}^n | \tilde{U}^n = U_1^n, M_1=1}, p_{\tilde{Z}_1^n | \tilde{U}^n = U_1^n} \right) \right] &\leq 4\tau + \frac{4}{\gamma_n^2} \mathbb{P}_{\tilde{U}^n \tilde{X}^n \tilde{Z}^n} \left[\frac{1}{n} \mathbb{I}(\tilde{X}^n; \tilde{Z}^n | \tilde{U}^n) \geq \frac{\log M'_1}{n} + \frac{\log \rho}{n} \right] \\ &\quad + \frac{4}{\gamma_n^2} \mathbb{P}_{\tilde{U}^n \tilde{X}^n \tilde{Z}^n} \left[\frac{1}{n} \mathbb{I}(\tilde{X}^n; \tilde{Z}^n | \tilde{U}^n) \geq \frac{\log M'_1}{n} \right] + \frac{4 \cdot 2^{-n\gamma}}{(\gamma_n \rho + 1 - \gamma_n)^2} \\ &\quad + \frac{4}{(\gamma_n \rho + 1 - \gamma_n)^2} \mathbb{P}_{\tilde{U}^n \tilde{X}^n \tilde{Z}^n} \left[\frac{1}{n} \mathbb{I}(\tilde{X}^n; \tilde{Z}^n | \tilde{U}^n) \geq \frac{\log M'_1}{n} - \gamma \right]. \quad (49) \end{aligned}$$

If $\frac{1}{n} \log M'_1 \geq \mathbb{I}(\tilde{X}; \tilde{Z} | \tilde{U}) + 2\gamma$, then Lemma 7 guarantees there exists $\alpha_\gamma > 0$ such that

$$\mathbb{P}_{\tilde{U}^n \tilde{X}^n \tilde{Z}^n} \left[\frac{1}{n} \mathbb{I}(\tilde{X}^n; \tilde{Z}^n | \tilde{U}^n) \geq \frac{\log M'_1}{n} - \gamma \right] \leq 2^{-\alpha_\gamma n}. \quad (50)$$

Set $\tau = 2^{-\eta n}$ for some η such that $0 < 2\eta < \min(\gamma, \alpha_\gamma)$; note that $\rho = \frac{\ln 2}{2} 2^{-\eta n} + o(2^{-\eta n})$. Therefore, for n large enough,

$$\frac{1}{n} \log \rho \geq -\gamma, \quad \frac{1}{\gamma_n \rho + 1 - \gamma_n} \leq 2 \cdot 2^{\eta n}, \quad \frac{1}{\gamma_n^2} \leq 2. \quad (51)$$

Consequently, combining (46), (47), (49), (50) and (51), we obtain for n large enough,

$$\mathbb{E}_{C_n} [\mathbb{S}_2(C_n)] \leq 4 \cdot 2^{-\eta n} + 8 \cdot 2^{-\alpha_\gamma n} + 8 \cdot 2^{-\alpha_\gamma n} + 16 \cdot 2^{-(\gamma-2\eta)n} + 16 \cdot 2^{-(\alpha_\gamma-2\eta)n} + 2 \cdot 2^{-\beta_\delta n}.$$

Therefore, for n large enough, there exists $\alpha_{\gamma,\delta} > 0$ such that $\mathbb{E}[\mathbb{S}_2(C_n)] \leq 2^{-\alpha_{\gamma,\delta} n}$. ■

Using Markov's inequality and for n sufficiently large, we conclude that if

$$\begin{aligned} R_0 &\leq \min \left(\mathbb{I}(\tilde{U}; \tilde{Y}) - 2\gamma, \mathbb{I}(\tilde{U}; \tilde{Z}) - 2\gamma \right) \\ R_1 &\leq \mathbb{I}(\tilde{X}; \tilde{Y} | \tilde{U}) - \mathbb{I}(\tilde{X}; \tilde{Z} | \tilde{U}) - 4\gamma, \end{aligned}$$

then there exists a specific code C_n such that $\mathbb{P}_e(C_n) \leq 2\epsilon$ and $\mathbb{S}_2(C_n) \leq 2^{-\frac{\alpha_{\gamma,\delta}}{2} n}$. Using [13, Lemma 1] with n large enough, we obtain $\mathbb{S}_1(C_n) \leq 2^{-\beta_{\gamma,\delta} n}$ for some $\beta_{\gamma,\delta} > 0$.

APPENDIX G

CONVERSE PART OF THEOREM 4

In the following, all equalities should be understood to hold with probability one. First, note that

$$\begin{aligned}
I(Y^n; X^n | Z^n) &\stackrel{(a)}{=} I(Y^n R_Y; X^n R_X | Z^n) \\
&= I(Y^n R_Y; X^n R_X A_1 | Z^n) - I(Y^n R_Y; A_1 | X^n R_X Z^n) \\
&\stackrel{(b)}{=} I(Y^n R_Y; A_1 | Z^n) + I(Y^n R_Y; X^n R_X | A_1 Z^n) \\
&= I(Y^n R_Y; A_1 | Z^n) + I(Y^n R_Y B_1; X^n R_X | A_1 Z^n) - I(B_1; X^n R_X | A_1 Y^n R_Y Z^n) \\
&\stackrel{(c)}{=} I(Y^n R_Y; A_1 | Z^n) + I(B_1; X^n R_X | A_1 Z^n) + I(X^n R_X; Y^n R_Y | Z^n A_1 B_1),
\end{aligned} \tag{52}$$

where (a) follows from the independence between R_X , R_Y and the source, (b) follows from $I(Y^n R_Y; A_1 | X^n R_X Z^n) = 0$ and (c) follows from $I(B_1; X^n R_X | A_1 Y^n R_Y Z^n) = 0$. By induction, we obtain

$$I(X^n; Y^n | Z^n) = I(X^n R_X; Y^n R_Y | Z^n A^r B^r) + \sum_{i=1}^r I(Y^n R_Y; A_i | Z^n A^{i-1} B^{i-1}) + \sum_{i=1}^r I(X^n R_X; B_i | Z^n B^{i-1} A^i).$$

Next, notice that

$$\begin{aligned}
I(X^n R_X; Y^n R_Y | Z^n A^r B^r) &= I(X^n R_X K; Y^n R_Y | Z^n A^r B^r) - I(K; Y^n R_Y | Z^n X^n R_X A^r B^r) \\
&\stackrel{(a)}{=} I(K; Y^n R_Y | Z^n A^r B^r) + I(X^n R_X; Y^n R_Y | K Z^n A^r B^r) \\
&= I(K; Y^n R_Y \hat{K} | Z^n A^r B^r) - I(K; \hat{K} | Z^n Y^n R_Y A^r B^r) + I(X^n R_X; Y^n R_Y | K Z^n A^r B^r) \\
&\stackrel{(b)}{=} I(K; \hat{K} | Z^n A^r B^r) + I(K; Y^n R_Y | Z^n A^r B^r \hat{K}) + I(X^n R_X; Y^n R_Y | K Z^n A^r B^r),
\end{aligned} \tag{54}$$

where (a) follows from $I(K; Y^n R_Y | Z^n X^n R_X A^r B^r) = 0$ and (b) follows from $I(K; \hat{K} | Z^n Y^n R_Y A^r B^r) = 0$.

Finally,

$$\begin{aligned}
I(K; \hat{K} | Z^n A^r B^r) &= I(K; \hat{K} Z^n A^r B^r) - I(K; A^r B^r Z^n) \\
&= H(K) - H(K | \hat{K} A^r B^r Z^n) - I(K; A^r B^r Z^n).
\end{aligned} \tag{55}$$

Combining (53), (54) and (55), we obtain

$$\begin{aligned}
H(K) &= I(X^n; Y^n | Z^n) + H(K | \hat{K} A^r B^r Z^n) + I(K; A^r B^r Z^n) - I(K; Y^n R_Y | Z^n A^r B^r \hat{K}) \\
&\quad - I(X^n R_X; Y^n R_Y | K Z^n A^r B^r) - \sum_{i=1}^k I(Y^n R_Y; A_i | Z^n A^{i-1} B^{i-1}) - \sum_{i=1}^k I(X^n R_X; B_i | Z^n B^{i-1} A^i)
\end{aligned}$$

and, consequently,

$$\begin{aligned} \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(K) &\leq \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y^n | Z^n) + \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} H(K | \hat{K} A^r B^r Z^n) + \text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(K; A^r B^r Z^n) \\ &\quad - \text{p-liminf}_{n \rightarrow \infty} I(K; Y^n R_Y | Z^n A^r B^r \hat{K}) - \text{p-liminf}_{n \rightarrow \infty} I(X^n R_X; Y^n R_Y | K Z^n A^r B^r) \\ &\quad - \sum_{i=1}^r \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(Y^n; A_i | Z^n A^{i-1} B^{i-1}) - \sum_{i=1}^r \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(X^n; B_i | Z^n B^{i-1} A^i) \end{aligned}$$

By assumption, $\lim_{n \rightarrow \infty} \mathbb{S}_6(\mathcal{S}_n) = 0$, hence $\text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} I(K; A^r B^r Z^n) = 0$. Similarly, since $\lim_{n \rightarrow \infty} \mathbb{P}_e(\mathcal{S}_n) = 0$, the Verdú-Han Lemma ensures $\text{p-limsup}_{n \rightarrow \infty} \frac{1}{n} H(K | \hat{K} A^r B^r Z^n) = 0$ and, since $\lim_{n \rightarrow \infty} \mathbb{U}(\mathcal{S}_n) = 0$, $\text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} H(K) = R_k$. In addition, note that

$$\begin{aligned} \forall i \in [1, r] \quad \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(Y^n R_Y; A_i | Z^n A^{i-1} B^{i-1}) &\geq 0 \quad \text{and} \quad \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(X^n R_X; B_i | Z^n B^{i-1} A^i) \geq 0 \\ \text{p-liminf}_{n \rightarrow \infty} I(K; Y^n R_Y | Z^n A^r B^r \hat{K}) &\geq 0 \quad \text{and} \quad \text{p-liminf}_{n \rightarrow \infty} I(X^n R_X; Y^n R_Y | K Z^n A^r B^r). \end{aligned}$$

Therefore,

$$R_k \leq \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y^n | Z^n).$$

Repeating the same argument starting from $I(X^n; Y^n)$ in place of $I(X^n; Y^n | Z^n)$, we obtain

$$R_k \leq \text{p-liminf}_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y^n).$$

REFERENCES

- [1] A. D. Wyner, "The Wire-Tap Channel," *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1367, October 1975.
- [2] I. Csiszár and J. Körner, "Broadcast Channels with Confidential Messages," *IEEE Trans. on Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.
- [3] Y. Liang, H. V. Poor, and S. S. (Shitz), *Information-Theoretic Security*, ser. Foundations and Trends in Communications and Information Theory. Delft, Netherlands: Now Publishers, 2009, vol. 5, no. 1–5.
- [4] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, October 2011.
- [5] T. Han and S. Verdú, "Approximation Theory of Output Statistics," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 752–772, May 1993.
- [6] J. G. D. Forney, "On the Role of MMSE Estimation in Approaching the Information-Theoretic Limits of Linear Gaussian Channels: Shannon Meets Wiener," in *Proc. of 41st Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, October 2003, pp. 430–439.
- [7] T. S. Han, *Information-Spectrum Methods in Information Theory*. Springer, 2002.
- [8] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian Wire-Tap Channel," *IEEE Trans. Inf. Theory*, vol. 24, no. 4, pp. 451–456, July 1978.

- [9] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Secure Communication Over Fading Channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2470–2492, June 2008.
- [10] U. Maurer, "Secret Key Agreement by Public Discussion from Common Information," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.
- [11] R. Ahlswede and I. Csiszár, "Common Randomness in Information Theory and Cryptography. I. Secret Sharing," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1121–1132, July 1993.
- [12] U. M. Maurer and S. Wolf, "Information-Theoretic Key Agreement: From Weak to Strong Secrecy for Free," in *Advances in Cryptology - Eurocrypt 2000*, Lecture Notes in Computer Science. B. Preneel, 2000, p. 351.
- [13] I. Csiszár, "Almost Independence and Secrecy Capacity," *Problems of Information Transmission*, vol. 32, no. 1, pp. 40–47, January–March 1996.
- [14] C. H. Bennett, G. Brassard, C. Crépeau, and U. Maurer, "Generalized Privacy Amplification," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1915–1923, November 1995.
- [15] M. Hayashi, "General Nonasymptotic and Asymptotic Formulas in Channel Resolvability and Identification Capacity and their Application to the Wiretap Channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1562–1575, April 2006.
- [16] I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 44–55, January 2005.
- [17] S. Verdú and T. S. Han, "A General Formula for Channel Capacity," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1147–1157, July 1994.
- [18] H. Koga, "Coding Theorems on Shannon's Cipher System with a General Sources," in *Proc. IEEE International Symposium on Information Theory*, Sorrento, Italy, June 2000, pp. 158–.
- [19] H. Koga and N. Sato, "On an Upper Bound of the Secrecy Capacity for a General Wiretap Channels," in *Proc. International Symposium on Information Theory*, Adelaide, Australia, September 2005, pp. 1641–1645.
- [20] D. Kobayashi, H. Yamamoto, and T. Ogawa, "How to Attain the Ordinary Channel Capacity Securely in Wiretap Channels," in *Proc. IEEE Information Theory Workshop on Theory and Practice in Information-Theoretic Security*, Awaji Island, Japan, October 2005, pp. 13–18.
- [21] M. Bloch and J. N. Laneman, "On the Secrecy Capacity of Arbitrary Wiretap Channels," in *Proceedings of 46th Allerton Conference on Communication, Control, and Computing*, Monticello, IL, September 2008, pp. 818–825.
- [22] M. R. Bloch, "Achieving Secrecy: Capacity vs. Resolvability," in *Proc. of IEEE International Symposium on Information Theory*, Saint Petersburg, Russia, August 2011, pp. 632–636.
- [23] J. Barros and M. Bloch, "Strong Secrecy for Wireless Channels," in *Information Theoretic Security*, ser. Lecture Notes in Computer Science. Calgary, Canada: Springer Berlin / Heidelberg, August 2008, pp. 40–53, (invited).
- [24] Y. Liang, G. Kramer, H. V. Poor, and S. S. (Shitz), "Compound Wiretap Channels," *EURASIP Journal on Wireless Communications and Networking*, vol. 142374, pp. 1–12, 2009.
- [25] T. Liu, V. Prabhakaran, and S. Vishwanath, "The Secrecy Capacity of a Class of Parallel Gaussian Compound Wiretap Channels," in *Proc. IEEE International Symposium on Information Theory*, July 2008, pp. 116–120.
- [26] M. Bloch, "Channel Intrinsic Randomness," in *Proc. of IEEE International Symposium on Information Theory*, Austin, TX, June 2010, pp. 2607–2611.
- [27] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes*. Holden Day, 1964.
- [28] S. Vembu and S. Verdú, "Generating Random Bits from an Arbitrary Source: Fundamental Limits," *IEEE Trans. Inf. Theory*, vol. 41, no. 5, pp. 1322–1332, September 1995.

- [29] M. Hayashi, "Exponential Decreasing Rate of Leaked Information in Universal Random Privacy Amplification," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3989–4001, 2011.
- [30] A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J.-M. Merolla, "Applications of LDPC Codes to the Wiretap Channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 8, pp. 2933–2945, August 2007.
- [31] H. MahdaviFar and A. Vardy, "Achieving the Secrecy Capacity of Wiretap Channels Using Polar Codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6428–6443, 2011.
- [32] L. Luzzi and M. R. Bloch, "Capacity-based random codes cannot achieve strong secrecy over symmetric wiretap channels," in *Proc. of 1st. International Workshop on Secure Wireless Networks*, Cachan, France, May 2011, (invited).
- [33] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel Coding Rate in the Finite Blocklength Regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.
- [34] R. Matsumoto and M. Hayashi, "Strong security and separated code constructions for the broadcast channel with confidential messages," October 2010. [Online]. Available: arXiv:1010.0743
- [35] M. Bloch and J. N. Laneman, "Information-Spectrum Methods for Information-Theoretic Security," in *Proc. Information Theory and Applications Workshop*, San Diego, CA, February 2009, pp. 23–28, (invited).
- [36] B. P. Dunn, M. Bloch, and J. N. Laneman, "Secure bits through queues," in *Proc. IEEE Information Theory Workshop on Networking and Information Theory*, Volos, Greece, June 2009, pp. 37–41.
- [37] Y. Sankarasubramanian, A. Thangaraj, and K. Viswanathan, "Finite-state wiretap channels: Secrecy under memory constraints," in *Proc. IEEE Information Theory Workshop*, October 2009, pp. 115–119.
- [38] A. J. Pierrot and M. R. Bloch, "Strongly Secure Communications Over the Two-Way Wiretap Channel," *IEEE Trans. Inf. Forensics and Security*, vol. 6, no. 3, pp. 595–605, September 2011.
- [39] A. Subramanian, A. Thangaraj, M. Bloch, and S. McLaughlin, "Strong Secrecy on the Binary Erasure Wiretap Channel Using Large-Girth LDPC Codes," *IEEE Trans. Inf. Forensics and Security*, vol. 6, no. 3, pp. 585–594, September 2011.
- [40] H. Yamamoto, "Rate-Distortion Theory for the Shannon Cipher Systems," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 827–835, May 1997.
- [41] S. Watanabe, T. Saitou, R. Matsumoto, and T. Uyematsu, "Strongly Secure Privacy Amplification Cannot be Obtained by Encoder of Slepian-Wolf Codes," in *Proc. IEEE International Symposium on Information Theory*, Seoul, Korea, July 2009, pp. 1298–1302.
- [42] P. W. Cuff, "Communication in Networks for Coordinating Behavior," Ph.D. dissertation, Princeton University, July 2009.
- [43] M. Hayashi, "Second-Order Asymptotics in Fixed-Length Source Coding and Intrinsic Randomness," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4619–4637, October 2008.
- [44] A. N. Shiryaev, *Probability*, 2nd ed. Springer, 1995.
- [45] X. He and A. Yener, "MIMO Wiretap Channels with Arbitrarily Varying Eavesdropper Channel States," submitted to *IEEE Trans. Inf. Theory*, September 2010.